DEBI-NN: Distance-Encoding Biomorphic-Informational Neural Networks for Minimizing the Number of Trainable Parameters

Laszlo Papp (✉ laszlo.papp@meduniwien.ac.at)
Medical University of Vienna  https://orcid.org/0000-0002-9049-9989

David Haberl
Boglarka Ecsedi
Clemens Spielvogel
Denis Krajnc
Marko Grahovac
Sasan Moradi
Wolfgang Drexler

Research Article

Keywords: biomorphic computing, informational neural network, distance encoding, machine learning, artificial intelligence, medical AI

Posted Date: December 20th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2384764/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at Neural Networks on August 25th, 2023. See the published version at https://doi.org/10.1016/j.neunet.2023.08.026.
DEBI-NN: Distance-Encoding Biomorphic-Informational Neural Networks for Minimizing the Number of Trainable Parameters

L. Papp¹, D. Haberl², B. Ecsedi¹, C. Spielvogel², D. Krajnc¹, M. Grahovac², S. Moradi¹, W. Drexler¹

1: Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria
2: Division of Nuclear Medicine, Medical University of Vienna, Vienna, Austria
3: Georgia Institute of Technology, Atlanta, GA, USA

*Corresponding author:
Laszlo Papp
Laszlo.papp@meduniwien.ac.at

Applied Quantum Computing Group, Center for Medical Physics and Biomedical Engineering
Medical University of Vienna
Waehringer Guertel 18-20
1090 Vienna, Austria

Keywords: biomorphic computing, informational neural network, distance encoding, machine learning, artificial intelligence, medical AI
Abstract

Modern artificial intelligence (AI) approaches mainly rely on neural network (NN) or deep NN methodologies. However, these approaches require large amounts of data to train, given, that the number of their trainable parameters has a polynomial relationship to their neuron counts. This property renders deep NN not applicable in fields operating with small, albeit representative datasets such as healthcare. In this paper, we propose a novel neural network architecture which trains spatial positions of neural soma and axon pairs, where weights are calculated by axon-soma distances of connected neurons. We refer to this method as distance-encoding biomorphic-informational (DEBI) neural network. This concept significantly minimizes the number of trainable parameters compared to conventional neural networks. We demonstrate that DEBI models can yield comparable predictive performance in tabular and imaging datasets, where they require a fraction of trainable parameters compared to conventional NNs, resulting in a highly scalable solution.

1. Introduction

With the advent of modern life’s digitalization endeavors, efforts to collect, process, and interpret large amounts of data have been actively conducted in various fields of research relying on artificial intelligence (AI) approaches (Bahrammirzaee, 2010; Hessler & Baringhaus, 2018; Li et al., 2017; K.-H. Yu et al., 2018). Neural networks (NN), in particular, have been in the forefront of AI research, as they are generally considered superior to many traditional machine learning approaches (Suzuki, 2017). The general concept of NNs is over half a century old, originally inspired by how the biological brain works (Rosenblatt, 1958). As such, most popular, so-called informational NNs follow the training concept to optimize weights between connected neurons (Akay et al., 2022). In case of deep neural networks – especially within fully-connected layers – this results in a polynomial increase of the number of the weights to train in relation to the number of neurons in the network. In contrast, so-called biomorphic NNs do not operate with learning weights, but with neurons composed of dendrites, somas, and axons to model neurophysiological processes of the biological brain (Filippov et al., 2020). In general, various informational and biomorphic NN concepts exist that model biological processes with different depths of complexities (Filippov et al., 2020). However, the vast majority of NN applications follow the informational approach (Schuman et al., 2017).

To date, various deep NN architectures have been proposed to successfully deal with complex problem domains such as image analysis (Ajit et al., 2020) or natural language processing (Young et al., 2018). Nevertheless, highly complex informational NNs require large amounts
of data and a relatively high number of parameters to train (Marcus, 2018), making their training complex and dependent on massive computational power (Anthony et al., 2020). In general, an AI model shall have more training samples than its number of trainable parameters, however, this alone does not guarantee avoiding overfitting (Belkin et al., 2019). Consequently, NN training approaches utilize various steps to minimize the chance of overfitting by e.g., epoch training, reduced model complexity and regularization (Wu et al., 2019). Nevertheless, even if all these methods are combined together to build NN models, they still tend to have low generalization abilities when applied to small training data (Borisov et al., 2021). Due to the above reasons, any research field without possessing the necessary amount of data, finds the exploitation of NNs challenging.

Pioneers of AI had recognized the above phenomenon and have been actively proposing that AI models need to undergo a major consolidation process in order to make them applicable to cases with small, albeit representative data (Motamedi et al., 2021; Zhong et al., 2022). Healthcare, in particular, is one of those fields, which struggles to deliver high-quality, curated big data, while it is also a prime candidate to benefit from the utilization of AI in order to support the advancement of personalized medicine (Litjens et al., 2017; Varoquaux & Cheplygina, 2022).

In light of the above, we propose a neural network architecture which considers neurons as soma-axon pairs that have three-dimensional spatial coordinates. Consistently, instead of training weights between connected neurons, we train spatial positions of soma-axon pairs, where the distance between connected neurons determines their weights. With our distance-to-weight mapping concept, we create a bridge between biomorphic and informational NNs. This approach allows to dramatically decrease the number of trainable parameters (i.e., the spatial coordinates of neurons) during the training process, since it maintains a linear relationship between the number of trainable parameters and the number of neurons. We refer to this concept as the Distance-Encoding Biomorphic-Informational Neural Network (DEBI-NN). Since a trained DEBI-NN model also operates with weights calculated from the distances between its neurons, it can be considered an informational NN.

We hypothesize that the DEBI-NN is capable to provide comparable performance to that of conventional informational NNs using small datasets, while keeping trainable parameter counts significantly lower. In order to test our hypothesis, we aimed to demonstrate DEBI-NN predictive performance in a supervised classification setting, hence, defined the following objectives: (a) to collect clinically-relevant tabular and imaging datasets with binary reference labels to predict; (b) to build binary classifier models relying on both conventional NN and
DEBI-NN model schemes with harmonized parameter sets, and (c) to compare validation predictive performances of NN and DEBI-NN models in the collected datasets.

2. Methods

2.1 Distance-Encoding Biomorphic-Informational Neural Network (DEBI-NN)

Neurons in DEBI-NNs resemble properties of biomorphic artificial neurons (Filippov et al., 2020), as they are composed of soma-axon terminal (axon) pairs, however, contrary to biomorphic models, the DEBI-NN neuron does not consider signal duration. Furthermore, in DEBI-NNs, each soma and axon has its own 3D (x,y,z) coordinates to train (Fig 1). This representation was chosen to eliminate the dependence of spatial alignments in-between consecutive, connected neural network layers. Due to this approach, the distance between a soma and the axons in the previous layer have no effect on the distance between the given neuron’s axon and the somas in the next layer. While the DEBI-NN scheme does not explicitly define a dendrite object type, each soma refers to previous layer connected neurons as “dendrite”. Consistently, the distance of a neuron and its dendrites denotes the distance of the neuron’s soma and the axons of its dendrites. Furthermore, similarly to biological brains, the signal between the soma and the axon of a neuron is not deteriorating. For details of how a DEBI-NN coordinate system is handled during training see Sec. 2.2.

![Diagram of DEBI-NN and Informational NN](image-url)
Figure 1: The principles of informational (left) vs. distance-encoding biomorphic-informational (DEBI) (right) artificial neurons with two consecutively connected neural network layers. Informational artificial neurons calculate their output by weighted averaging their inputs from previous layer neurons where weights ($w_i$) are trained. In contrast, DEBI neurons train their soma ($s_i$) and axon ($a_i$) spatial coordinates to calculate distances ($d_i$) in-between previous layer axons and consecutive layer somas. The weight ($w_i$) between connected DEBI neurons is calculated by their specific distance-to-weight ($f_D(…)$) functions (See Supplemental Appendix A). DEBI neural networks are unidirectional, meaning that soma-axon pairs as well as consecutive layers are aligned across a positive an increasing $z$-coordinate, while their $x$ and $y$ coordinates can both be positive or negative.

2.2 DEBI-NN coordinate system to train

A soma $x,y,z$ coordinate is relative to the spatial frame of reference of its given layer. In addition, the $z$-coordinate of an axon is also relative to the $z$-coordinate of the given neuron’s soma (Fig. 2). In the current DEBI-NN concept, any soma or axon $z$-coordinate can be only positive to ensure signal unidirectionality within the network. Due to the relative 3D coordinate system of spatial neurons, training their coordinate values results in a balanced parameter search space and does not inherently require training schemes such as gradient descent (Goodfellow et al., 2016). In order to eliminate any conventional NN-specific inherent bias, training a DEBI-NN is currently conceptualized to follow simple evolutionary training algorithms (S. Yang, 2007). However, the current DEBI-NN model scheme does not rule out the utilization of other training schemes. See Supplemental Appendix C for details of how DEBI-NNs were trained in this study. Note that weights are not only influenced by soma and axon distances or the distance-to-weight mapping function (See Supplemental Appendix A), but also by the overlap ratio of its consecutive spatial layers. For details see Sec. 2.3.

2.3 Spatially-overlapping DEBI-NN layers

Weight variations in DEBI-NNs can be influenced not only by the training process itself, but also by a modifiable overlap ratio of their spatial layers (Fig. 2). Note that while the layer overlap ratio can be set as a trainable parameter, this study relied on fixed overlap ratios across all the involved datasets (See Supplemental Appendix A for DEBI-NN and NN parameters). Increasing the spatial overlap in-between layers increases the fidelity of possible distances (and calculated weights) between connected neurons, while increasing the gap between consecutive layers makes distances more and more similar, thus, it flattens weight variations.
Figure 2: Two-dimensional representation of DEBI-NN neurons. A DEBI-NN neuron is composed of a soma (s) and an axon (a) 3D coordinate pair. Given two hidden layers (H_1 and H_2), they both have a spatial volume in 3D as well as their DEBI neurons inside. A DEBI layer’s spatial volume is the smallest bounding box, enclosing all its spatial neurons. This bounding box is aligned to the x,y,z coordinate system of the given network. Z-coordinates of somas are relative to their layer z-coordinate and axon z-coordinates are relative to their soma z-coordinates. During a DEBI-NN training process, soma and axon coordinates are trained. Somas collect input signals and generate their action potential (x_1, x_2 for hidden layer 1 neurons and y_1, y_2 for hidden layer 2 neurons) which is sent to their axons without deterioration. Axons further signal the action potential to neural somas in the next layer. In alignment with the concept in Fig. 1, distances (d_1,…,d_4) of DEBI neurons are determined as current layer soma and previous layer axon distances, from which weights are calculated (see Fig. 1). Note that in DEBI-NNs the z-overlap ratio of consecutive layers (e.g., H_1,r) can be a fixed value or a hyperparameter. Consecutive layer overlap ratios denote the ratio in-between the portion of the next layer’s z-width (e.g., H_2.zw) and the given layer’s z-width (e.g., H_1.zw) spatially overlapping. This example demonstrates the schematics of a DEBI-NN in 2D in order to simplify the view.

2.4 DEBI-NN vs. conventional NN compatibility
Backward compatibility of a DEBI-NN to conventional NNs is provided, given, that a spatially connected and trained DEBI-NN also operates with weights; with the difference that those
weights are not trained, but calculated from trained neural soma-axon distances. For accessing the detailed technical documentation and implementation of the DEBI-NN see Sec. 6.

2.5 Datasets
This study involved four open-access medical datasets for its analysis, including Mammographic mass (Elter et al., 2007) and Heart failure (Chicco & Jurman, 2020) tabular data as well as two imaging datasets including Breast MNIST (Al-Dhabyani et al., 2020) and Pneumonia MNIST (J. Yang, Shi, & Ni, 2021; J. Yang, Shi, Wei, et al., 2021). All experiments of this study were performed in accordance with the respective guidelines and regulations of the involved open-access data sources. For access of the datasets, see Sec. 6.

The collected datasets demonstrate a diverse pattern regarding sample (n=299-5856) and feature counts (n=5-196), feature-to-sample count ratios (FSR=0.5%-35.8%) as well as minority class imbalance ratios (26-46%). Tabular dataset feature counts are considered after performing feature ranking and selection (Sec. 2.7). MNIST imaging dataset feature counts are identical with the number of pixels involved in the analysis of each sample (Sec. 2.7). See Table 1 for dataset characteristics.

2.6 Trainable parameters and network configurations
In this study, the total number of trainable parameters of a conventional NN is given by the sum of the number of weights in-between connected neurons given by equation (1):

\[
P_{NN} = \sum_{i=1}^{H+1} N_i N_{i-1}
\]

(1)

where \( N_i \) is the number of neurons in the \( i^{th} \) layer (\( 1 \leq i \leq H + 1 \)) and \( H \) is the number of hidden layers in the given fully-connected NN.

In contrast, the total number of parameters in the DEBI-NN consists of 3D spatial coordinates of the neurons provided by equation (2):

\[
P_{DBNN} = 3I + 3O + \sum_{i=1}^{H} 6N_i
\]

(2)

Where \( I \) is the number of input neurons, \( O \) is the number of output neurons, \( N_i \) is the number of hidden neurons in the \( i^{th} \) hidden layer (\( 1 \leq i \leq H \)) in a fully-connected DEBI-NN. Note that input neurons have only axon 3D coordinates, while output neurons have only soma 3D
coordinates to train, as they are relevant for determining distances, and thus, weights in the network. In contrast, hidden neurons have 3+3 parameters to train for their somas and axons respectively (see Sec. 2.1).

This study operated with one hidden layer in both DEBI-NNs and NNs to conduct a fair comparison. See Supplemental Appendix A for detailed NN and DEBI-NN parameter sets. Parameter count properties of DEBI-NN and NN model schemes are provided in Table 2.

### 2.7 Cross-Validation and Data Preprocessing

Each tabular dataset underwent a random training-test subset split 100-times with a 80-20% split ratio respectively, considering the minority subset count, hence, resulting in a balanced validation subset for each fold (Papp et al., 2021). Consistently, the given training subset remained imbalanced. Z-score normalization across all features in the given tabular data training subset was performed (LeCun et al., 2012). The calculated mean-deviation values of the training subset for the z-score normalization served as inputs for normalizing the respective test subset as well in each of the 100 folds. Feature ranking and selection was performed in each tabular data training subset utilizing the Minimal-redundancy-maximal-relevance (mRMR) approach (Hanchuan Peng et al., 2005) which considers feature redundancy analysis as well. The respective test subset was also reduced to the selected features of the training subset. In case of tabular datasets, the training subset of each fold was further split to train-validate subsets during the training process for both DEBI-NNs and conventional NNs as of their own internal mechanisms. For further details, see Supplemental Appendix A.

MNIST imaging datasets were resampled to 14x14 pixels via local averaging, thus, resulting in overall 196 pixels per sample. Since the MNIST datasets were originally published with train-validate-test splits, their split configurations were used in an as-is basis.

Building models was performed on train subsets, while fitness measurements for early stopping were performed on the validate subset of the given model. Reported predictive performance values of this study were calculated for each model by evaluating their independent test subset samples that were not part of training or early stopping decision making processes.

Class imbalance was handled directly in the entropy loss calculation in both DEBI-NNs and conventional NNs which was utilized to assign fitness to model variants during the training process (Fernando & Tsokos, 2022).
2.8 Predictive performance estimation

Predictive performance estimations of DEBI-NNs and conventional NNs were evaluated relying on cross-validation and data preprocessing as presented in Sec. 2.7 in each dataset, predicting their reference binary labels (see Table 1). Dataset-specific DEBI-NN and NN models were built with the same network configuration (see Table 2). See Supplemental Appendix A for parameter sets of DEBI-NNs and NNs built in this study.

The number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) cases were calculated across each independent test subset samples of the given model. Mean sensitivity (SNS), specificity (SPC), positive predictive value (PPV), negative predictive value (NPV) and accuracy (ACC) were calculated for predictive performance comparison across samples of test subsets of each model.

3. Results

3.1 Network parameter count comparison

By increasing the number of neurons in the network, the neuron-to-trainable parameter ratio shows a polynomial growth in case of conventional NNs, however, the same trend remains linear in case of DEBI-NNs (Fig. 3). As shown in Table 2, DEBI-NNs have a significantly lower trainable parameter count to train even in case of small tabular datasets. For the demonstration of parameter counts in DEBI-NNs and conventional NNs with different numbers of hidden layers, see Fig. 3.
Figure 3: Relationship of neuron count and weight-related trainable parameter count in conventional (NN) vs. distance-encoding biomorphic-informational neural networks (DEBI-NN) with different hidden layer counts. These schemes perform binary classification with one (1H), two (2H) and three (3H) hidden layers, where hidden neuron count follows the rule as provided in (Heaton, 2015). Note that trainable parameters in case of NNs are weights, while in case of DEBI-NNs they are spatial coordinates of neurons in the DEBI-NN.

DEBI-NN vs. conventional NN parameter count break-even is achieved with networks having ~10 neurons, while significant differences in parameter counts are already present in relatively small networks. For instance, with n=300 neurons the trainable parameter count ratio in fully-connected DEBI-NNs vs. NNs is 6%, 9% and 10% with 1, 2 and 3 hidden layers respectively (Fig. 3). See Supplemental Appendix B for detailed parameter count estimations with different neuron count and hidden layer count configurations in NNs and DEBI-NNs.

3.2 DEBI-NN spatial properties during and after training

Figure 4 provides insights into the process of DEBI-NN training. The initial network is small and dense in volume and is iteratively growing in 3D to achieve a spatial distribution which increases its fidelity to build optimal DEBI-NN models (Fig. 4 A-D). This behavior is in line with the utilized distance-to-weight function, which relies on a Gaussian function having a fixed half-maximum (See Supplemental Appendix A).
Figure 4: Three-dimensional rendering of selected training phases (A-D) of a DEBI-NN in the Hearth failure dataset. Input, hidden and output neurons are colored by red, yellow and green respectively. Somas have twice the diameter than their respective axons. Soma-axon connections are colored by their respective neuron color. Weights are colored by shades of blue (light blue - low weights, dark blue - high weights). For a video animation of how a DEBI-NN model was trained in 3D see Sec. 6.

Figure 5 represents a trained network (Fig. 5, A) which moved input axons into a parabolic shape, where input axons corresponding to low-importance input features are organized towards the outer regions of the network in 3D (Fig. 5, B). In contrast, high-importance input axons are near and surrounded by various hidden neurons (Fig. 5, C). Furthermore, output neurons are organized in the middle of the network, surrounded by the hidden layer somas and axons to achieve a spatial configuration which maximizes the distance-weight fidelity of the network (Fig. 5, D). For a video of a DEBI-NN training animation, see Sec. 6.
Figure 5: Three-dimensional rendering of a trained Distance-Encoding Biomorphic-Informational Neural Network (A) in the Hearth failure dataset. Input, hidden and output neurons are colored by red, yellow and green respectively. Somas have twice the diameter than their respective axons. Soma-axon connections are colored by their respective neuron color. Weights are colored by shades of blue (light blue - low weights, dark blue - high weights). B: distant input axon, corresponding to a low importance input neuron; C: high weights in-between input axon and several hidden somas around; D: output neuron somas (green) are in the center of the network, while all connected axons from previous layers (yellow) were organized around them in 3D to optimize distance distributions. For a video animation of how a DEBI-NN model was trained in 3D see Sec. 6.

3.2 Predictive performance evaluation

As shown in Table 3, DEBI-NNs provided comparable predictive performance in all included datasets of this study. In both tabular datasets, DEBI-NNs resulted in comparable Confidence Interval (CI) percentage variations as well. Overall, DEBI-NN models resulted in more
balanced classifiers (see SNS-SPC and PPV-NPV pairs in Table 3) compared to conventional NNs in the function of increasing the number of training samples (see Table 1 and 3).

4. Discussion

In this study we proposed a distance-encoding biomorphic-informational neural network (DEBI-NN) architecture to minimize the number of trainable parameters compared to those of conventional neural networks (NN) while maintaining performance on small datasets. The DEBI-NN models yielded identical accuracies (ACC) compared to conventional NNs with substantially less trainable parameters (6.6 – 88%) across four different medical datasets. In general, DEBI-NNs demonstrated to lead to more balanced classifiers in terms of SNS-to-SPC and PPV-to-NVP ratios with the increase of the number of samples to train on, even in the presence of a high minority class imbalance (see Table 1, 2 and 3). We consider that this phenomenon is due to the significant parameter count decrease in DEBI-NNs, which supported the training process to converge to an optimal balanced entropy loss.

Our findings are relevant for clinical scenarios, where both sensitivity and specificity (as well as positive and negative predictive values) need to be maximized and preferably balanced to minimize the chances of misclassifications. This is particularly relevant in healthcare, as most disease subtypes are naturally imbalanced (Krajnc et al., 2021). The ability to build balanced classifiers with DEBI-NNs while maintaining a significantly lower number of trainable parameters compared to NNs has profound implications, as this property allows to build DEBI-NN classifiers on relatively small, but representative datasets in the future. Furthermore, since DEBI-NNs are 3D spatial networks, their interactive visualization supports interpretation processes that are crucial, especially in fields such as healthcare (Gemson Andrew Ebenezer & Durga, 2015; Papp et al., 2018). This added value renders DEBI-NNs potential candidates to support the building and clinical adoption of medical AI applications that are to date, underrepresented (Bradshaw et al., 2022).

The added value of incorporating spatial information to informational neural networks has been investigated by Wołczyk et al. (Wołczyk et al., 2019). In their model scheme, neurons belonging to the same layer were placed on a 2D plane, where neuron-neuron distance within the given layer was utilized as regularization factor to cluster neurons belonging to the same task to perform. Unlike our proposed architecture, the concept of Wołczyk et al. (Wołczyk et al., 2019) does not distinguish somas and axons, and does not rely on in-between layer neuron distances to determine weights. Furthermore, while graph neural networks (GNNs) (Scarselli et al., 2009) may resemble properties of encoding rich information among data elements (e.g.,
distances), they do not consider training spatial positions of their neurons. In contrast, our model scheme represents spatial relationships within the network through its unique architecture to model any kind of data. In addition to informational (a.k.a. weights are trained) neural networks, various biomorphic neural networks have been proposed in the literature to model biological processes of the human brain with different magnitudes of complexities (Schuman et al., 2017). In this regard, computational models have been created that aim to bridge the efficiency of conventional NNs and the favorable properties of real neural network circuits, such as analog computation, low power consumption or fast inference (Pfeiffer & Pfeil, 2018). The most successful implementations include spiking neural networks (SNNs) (Pfeiffer & Pfeil, 2018) and deep spiking neural networks (DSNNs). SNNs follow the Hebbian learning pattern, meaning that whenever two connected neurons fire at the same time, the connection between them is strengthened. Despite their desirable properties that mimic those of real neural circuits, SNNs have the major disadvantage of low accuracy compared to their informational machine learning counterparts (Pfeiffer & Pfeil, 2018). Furthermore, while certain biomorphic networks conceptualize dendrites, somas and axons as individual network entities, they do not assign spatial properties to them (Filippov et al., 2020). Based on the corresponding literature, we consider the concept of DEBI-NN a hybrid solution that possesses features of both biomorphic and informational neural networks, as it acts like a spatial biomorphic NN during training, and results in a conventional informational NN once training is completed.

While our study analyzed medical datasets that are naturally small and imbalanced, we wish to emphasize that the DEBI-NN concept is generic, and hence, can have multiple implications in the field of NN research, as it could be used to model various network architectures to process a variety of different data types from images to sequential data. For instance, allowing bidirectionality of somas and axons in the DEBI-NN scheme could model recurrent neural networks (Y. Yu et al., 2019). Furthermore, convolutional neural networks (CNN) (Young et al., 2018) could be modeled by guided soma spatial positions relative to the input data or to the previous layer's axons. In theory, in case of a CNN operating with images, the DEBI-NN could generate not only rectangular, but circular or any irregular variations of convolutional kernels to better fit to the properties of the given input data (Ma et al., 2017). Transformers (Vaswani et al., 2017) could also be formulated by DEBI-NN neuron clusters that incorporate both close and distant parts of the input data at the same time. Beyond the above potential to utilize DEBI-NNs in various NN model schemes, we assume that the significantly lower parameter count of DEBI-NNs may support use cases operating with low-tier hardware and/or small training datasets, and in general, it can result in a lower economic and environmental footprint to train NNs (Dodge et al., 2022; Selvan et al., 2022). Furthermore, DEBI-NNs may
also allow to build highly-complex informational NN schemes such as GPT-3, Gopher or MT-NLG 530B (Hoffmann et al., 2022) with significantly fewer parameters. At the same time, DEBI-NNs may allow significantly more complex models to be built with the same number of parameters than the afore-mentioned large-scale models operate with. Nevertheless, investigations in this manner will require the current CPU-only DEBI-NN implementation to undergo an optimal GPU migration process.

While we recognize the future potentials of DEBI-NNs, this study also had notable limitations. As such, the demonstration of DEBI-NNs and NNs were performed in a binary classification setting and with a fixed parameter set across all included datasets which might have resulted in suboptimal predictive performance estimations. Nevertheless, we explicitly intended to demonstrate the concept of DEBI-NNs in a simple training setting, albeit in representative clinical datasets. Furthermore, given that the concept of DEBI-NNs is novel, to date, the list of its hyperparameters and their potential value ranges are unknown. In this regard, further investigations will need to be performed to determine the exact behavior, and thus, to reveal the true potential of DEBI-NNs in various parameter configurations and in different use cases. Further to the above, no regularization was performed in DEBI-NNs contrary to the utilized NNs for conducting the comparison analysis. The reason of this approach was that it is currently unclear whether regularizations routinely performed in conventional NNs are beneficial in DEBI-NNs, as NN regularizations are specific to how weights are trained. Hence, DEBI-NNs may either require regularization in their distance-to-weight functions, or – given that spatial configurations of a DEBI-NN affect all connected neurons – they may not require certain regularizations at all. In this regard, we currently hypothesize that a fully-connected DEBI-NN by design, may have self-regulating properties. Here, comparison studies shall be conducted investigating both generic and DEBI-NN-specific regularization approaches in order to explore how and if DEBI-NNs can be effectively regulated.

It is also important to emphasize that unlike traditional informational NNs, DEBI-NNs may not be able to result in weight configurations, which informational NNs – by training each weight individually – can do. Nevertheless, we consider various possible solutions in case spatial dependency in DEBI-NNs may prove suboptimal in specific use cases. First, beyond investigating regularization techniques in DEBI-NNs, the ability to significantly minimize the number of trainable parameters also allows to increase the number of hidden layers and neurons to counter-balance any potential predictive performance loss. Second, sparse DEBI-NNs may allow the formulation of networks to decrease spatial dependency which is currently originated from the fully-connected property of DEBI-NNs. Third, allowing DEBI neurons to have higher (>3) dimensional coordinates could provide additional degrees of freedom to further increase predictive performance. Last, DEBI-NNs could also be considered as pre-
trainers on lower number of trainable parameters (and/or lower number of training samples) before fine-tuning their individual weights by additional, conventional NNs as part of a specific DEBI-NN-to-NN transfer learning approach.

5. Conclusions

Our novel DEBI-NN concept trains spatial coordinates of somas and axons instead of weights in-between connected neurons. This property keeps the neuron-to-trainable parameter count relationship linear instead of polynomial, regardless of the size of the network. Our DEBI-NN models yielded similar predictive performance, while relying on significantly lower number of trainable parameters compared to conventional NNs in tabular and imaging datasets. Thus, we consider that DEBI-NNs have profound implications in the field of neural network-related research.

6. Data Availability

Open-access datasets involved in this study are found under the following links:

Mammographic mass (Elter et al., 2007):
http://archive.ics.uci.edu/ml/datasets/Mammographic+Mass

Heart failure (Chicco & Jurman, 2020):
https://archive.ics.uci.edu/ml/datasets/Heart+failure+clinical+records

Breast MNIST (Al-Dhabyani et al., 2020) and Pneumonia MNIST (J. Yang, Shi, & Ni, 2021; J. Yang, Shi, Wei, et al., 2021):
https://medmnist.com/

Source code repository for conventional ML evaluations and preprocessed data:
https://github.com/cspielvogel/DEBINN-Comparison

Doxygen documentation, DEBI-NN C++ implementation and pre-defined cohort executions including data and their settings utilized to conduct this study:
<[REDACTED TILL PEER REVIEW IS FINALIZED]>

3D animation video representing a DEBI-NN training process in the Heart Failure dataset:
https://youtu.be/S4Dj5qc7Rno
See Supplemental Appendix D for the utilized third-party software packages to conduct this study.

7. Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8. Acknowledgements

The authors would like to express their most sincere gratitude to John Cohn Ph.D. for his valuable insights and critical lecturing to finalize this paper.

9. Funding

This project was funded by the Medical University of Vienna (MUV), Center for Medical Physics and Biomedical Engineering. Computational calculations were provided by FocusXL-2019 grant “Foundations of a Quantum Computing Lab at the CMPBME” at the Medical University of Vienna as well as by the computational resources provided by the IT department (ITSC) of MUV.

10. Author Contributions

All authors contributed to writing and reviewing the paper. Specific contributions are as follows:

L. Papp: DEBI-NN concept, study design, DEBI-NN functional implementation, evaluation, literature research, repository management.

D. Haberl: conventional neural network (NN) evaluation, DEBI-NN and NN parameter harmonization, repository management, Linux code testing and evaluation.

B. Ecsedi: DEBI-NN rendering implementation, parameter count comparison, literature research.

C. Spielvogel: data preparation, conventional neural network (NN) evaluation, DEBI-NN and NN parameter harmonization, literature review.

D. Krajnc: data preparation, cross-validation scheme.

M. Grahovac: data collection, Doxygen documentation, hardware environment and computational resource management.

S. Moradi: set up evaluation hardware environment, porting source code to Linux.

W. Drexler: supervision, resources.
References


biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process. *Medical Physics*, 34(11), 4164–4172. https://doi.org/10.1118/1.2786864


Krajnc, D., Papp, L., Nakuz, T. S., Magometschnigg, H. F., Grahovac, M., Spielvogel, C. P.,

https://doi.org/10.1007/978-3-642-35289-8_3

intelligent manufacturing: a review. Frontiers of Information Technology & Electronic
Engineering, 18(1), 86–96. https://doi.org/10.1631/FITEE.1601885

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak,

http://arxiv.org/abs/1706.07966


Papp, L., Spielvogel, C. P., Grubmüller, B., Grahovac, M., Krajnc, D., Ecsedi, B., Sareshgi, R.
A. M., Mohamad, D., Hamboeck, M., Rausch, I., Mitterhauser, M., Wadsak, W., Haug, A.

Through Hybrid Imaging and Medical Big Data Analysis. Frontiers in Physics, 6. https://doi.org/10.3389/fphy.2018.00051


Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and


Table 1: Characteristics of the collected datasets. Sample and feature counts of datasets refer to the properties of the raw data without preprocessing or additional train-validate splitting. Minority class imbalance demonstrates the ratio of the minority subgroup size in relation to the total number of samples in the given training subset. For references to the original datasets, see Sec.2.5. FSR – Feature-to-sample count ratio; MNIST – Modified National Institute of Standards and Technology.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Mammographic mass</th>
<th>Heart failure</th>
<th>Breast MNIST</th>
<th>Pneumonia MNIST</th>
</tr>
</thead>
<tbody>
<tr>
<td># Samples</td>
<td>961</td>
<td>299</td>
<td>780</td>
<td>5856</td>
</tr>
<tr>
<td># Features</td>
<td>5</td>
<td>12</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>FSR</td>
<td>0.5%</td>
<td>4%</td>
<td>35.8%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Minority class imbalance</td>
<td>46%</td>
<td>32%</td>
<td>27%</td>
<td>26%</td>
</tr>
<tr>
<td>To predict</td>
<td>Malignancy (benign vs. malignant)</td>
<td>Died (yes vs. no)</td>
<td>Malignancy (normal or benign vs. malignant)</td>
<td>Pneumonia (yes vs. no)</td>
</tr>
<tr>
<td>Reference</td>
<td>(Elter et al., 2007)</td>
<td>(Chicco &amp; Jurman, 2020)</td>
<td>(Al-Dhabyani et al., 2020)</td>
<td>(Kermany et al., 2018)</td>
</tr>
</tbody>
</table>
Table 2: Basic parameters of distance-encoding biomorphic-informational (DEBI-NN) and conventional neural networks (NN) per dataset. Note that the parameter counts comparisons consider fully-connected neural networks and one hidden layer. For detailed NN and DEBI-NN parameter sets for each evaluation see Supplemental Appendix A. $P_{\text{NN}}$ – NN parameter count; $P_{\text{DEBI-NN}}$ – DEBI-NN parameter count; $R_P$ – DEBI-NN vs. NN parameter count ratio (%); MNIST – Modified National Institute of Standards and Technology.

<table>
<thead>
<tr>
<th></th>
<th>Mammographic mass</th>
<th>Heart failure</th>
<th>Breast MNIST</th>
<th>Pneumonia MNIST</th>
</tr>
</thead>
<tbody>
<tr>
<td># Input neurons</td>
<td>5</td>
<td>12</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td># Hidden Neurons</td>
<td>100</td>
<td>100</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td># Output Neurons</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$P_{\text{NN}}$</td>
<td>700</td>
<td>1400</td>
<td>99000</td>
<td>99000</td>
</tr>
<tr>
<td>$P_{\text{DEBI-NN}}$</td>
<td>621</td>
<td>642</td>
<td>6594</td>
<td>6594</td>
</tr>
<tr>
<td>$R_P$</td>
<td>88%</td>
<td>46%</td>
<td>6.6%</td>
<td>6.6%</td>
</tr>
</tbody>
</table>
Table 3: Cross-validation performance values of distance-encoding biomorphic-informational (DEBI-NN) and conventional neural networks (NN) per dataset. Performance values are presented in mean ratios ± the percentage of their confidence interval (CI) with 95% confidence level. MNIST imaging datasets have no CI, as they were published with one predefined train-validate-test subset configuration. SNS – Sensitivity; SPC – Specificity; PPV – Positive predictive value; NPV – Negative predictive value; ACC – Accuracy; MNIST – Modified National Institute of Standards and Technology.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Model</th>
<th>SNS</th>
<th>SPC</th>
<th>PPV</th>
<th>NPV</th>
<th>ACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammographic mass</td>
<td>DEBI-NN</td>
<td>0.85 ± 0.8%</td>
<td>0.82 ± 0.8%</td>
<td>0.82 ± 0.6%</td>
<td>0.84 ± 0.7%</td>
<td>0.83 ± 0.4%</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>0.82 ± 0.9%</td>
<td>0.83 ± 0.9%</td>
<td>0.83 ± 0.7%</td>
<td>0.82 ± 0.8%</td>
<td>0.83 ± 0.6%</td>
</tr>
<tr>
<td>Heart failure</td>
<td>DEBI-NN</td>
<td>0.81 ± 2.2%</td>
<td>0.78 ± 3.0%</td>
<td>0.80 ± 2.2%</td>
<td>0.81 ± 1.9%</td>
<td>0.79 ± 1.7%</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>0.75 ± 2.4%</td>
<td>0.83 ± 2.1%</td>
<td>0.82 ± 2.0%</td>
<td>0.77 ± 1.8%</td>
<td>0.79 ± 1.6%</td>
</tr>
<tr>
<td>Breast MNIST</td>
<td>DEBI-NN</td>
<td>0.81</td>
<td>0.67</td>
<td>0.87</td>
<td>0.56</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>0.77</td>
<td>0.76</td>
<td>0.90</td>
<td>0.55</td>
<td>0.77</td>
</tr>
<tr>
<td>Pneumonia MNIST</td>
<td>DEBI-NN</td>
<td>0.90</td>
<td>0.82</td>
<td>0.90</td>
<td>0.83</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>0.95</td>
<td>0.74</td>
<td>0.86</td>
<td>0.91</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PappetalDEBINNSupplemental.pdf