

Original Article

DNA-Pattern-Inspired Machine Learning Weight Optimisation

*Naveen Shrivastava¹, Bose Babu Ravula², Yudhveer Kandukuri³, Sojy John⁴, Sai Mahesh Marri⁵

¹T-Mobile Salesforce Lead Developer, USA.

²John Wiley & Sons, Inc., Principal, USA.

³G2 Web Services - Open Source DBA & DevOps, USA.

⁴Recreational Equipment, Inc - Technology Lead, DevOps Engineer & Cloud Migration, USA.

⁵Amazon Web Services, Senior Electrical Engineer, USA.

Abstract:

Combining molecular biology with machine learning has given rise to a new paradigm in which structural and functional characteristics of deoxyribonucleic acid (DNA) are used as design principles for optimising the weights of neural networks. This paper combines improvements from the last decade reported until 2020 in the fields of DNA-inspired weight encoding, enzymatic update rules, strand-displacement computation and evolutionary neural architecture search to provide a holistic framework, namely DNA-Pattern Inspired Machine Learning Weight Optimisation (DPIWMO). The results show that enzymatic weight update schemes [2] outperform conventional genetic algorithm (GA) based schemes [13] by 25.7% in terms of the number of parameters and 98.1% in terms of accuracy on benchmark logic problems. Winner-take-all (WTA) molecular networks with a 100-class classification [3] illustrate that the DNA substrate provides a search space of more than 10^{17} candidate weight configurations. It could be demonstrated that DNA computing-inspired deep architecture search [17] reached a top-1 accuracy of 67.4% on ImageNet, 3.2% better than the NEAT-based methods [10] which had only 16.4 million parameters. The convergence of network weights was improved by 18% and the training time was shortened by 2.6 hours for the Triplet codon encoding GAs than for binary encoded GAs [16]. The theoretical limit is explored, the No-Free-Lunch (NFL) theorem [14], which inspires use of hybrid approaches that integrate DNA-sequence motifs with gradient-based refinement. Cross domain utility is also demonstrated in applications for genomic sequence mining [15] which reaches 97.1% accuracy on 200,000 sequences. The proposed framework can be readily extended to the design of automated neural networks that are biologically plausible.

Keywords:

DNA Computing, Weight Optimization, Neural Architecture Search, Genetic Algorithms, Enzymatic Learning, Triplet Codon Encoding, Molecular Neural Networks, Evolutionary Computation, Deep Learning, Biologically Inspired Optimization, Strand Displacement, Particle Swarm Optimization.

Article History:

Received: 07.04.2022

Revised: 21.04.2022

Accepted: 30.04.2022

Published: 17.05.2022

1. Introduction

Optimisation of weights of an artificial neural network (ANN) is one of the basic problems in machine learning. However, gradient-based methods are easy to compute when the objective is differentiable, but can be trapped in local minima and are not appropriate for



the search space being discrete or non-differentiable. These restrictions inspire the search for alternatives inspired by nature, one of them being DNA-based computation, which is unique due to its high parallelism, high information density and well-defined molecular operations.

The concept of molecular computation was laid out by Adleman in 1994 when he demonstrated that DNA strands could encode and solve the Hamiltonian path problem [1]. In that seminal research, a directed graph with seven nodes was set into a strand of oligonucleotides, followed by an enzymatic ligation to count the number of valid paths—so effectively taking advantage of the 3×10^{23} molecules in a single micro-mole to perform a massively parallel combinatorial search. This proof of concept led to the next 20 years of research on the possibility of transfer of such encoding strategies to the field of neural network weight optimisation.

In 2011, Qian, Winfree and Bruck showed that threshold and supra-linear signal-restoration functions can be achieved with DNA strand displacement (DSD) cascades, and developed a network that could classify 4-bit binary patterns [9]. Cherry and Qian [3] followed with a 100-class winner-take-all (WTA) implementation of the DSD paradigm in 2018, where 1,800 molecular weight species were encoded into 50 microlitres with a pattern-recognition accuracy of 90.3% on a subset of hand-written digits, confirming the feasibility of large-scale molecular weight encoding.

Complementary algorithmic tools came in the form of parallel progress in evolutionary computation. Whitley [13] saw the canonical genetic algorithm (GA) as a population-based search procedure, in which the operators of crossover, mutation and selection are used to improve candidate weight vectors, represented as binary strings, over iterations. In 1995, Kennedy and Eberhart [6] introduced particle swarm optimisation (PSO), which was a model of collective intelligence, based on velocity updates from individual and neighbourhood best positions. In 2002, Stanley and Miikkulainen [10] introduced NeuroEvolution of Augmenting Topologies (NEAT), which allows for the simultaneous optimisation of network topology and weights using speciation and historical markings.

The amalgamation of these molecular and computational streams has led to a new class of algorithms that map onto the conceptual bases of DNA structural features (the four nucleotide alphabet, base-pairing complementarity, triplet codon semantics and enzymatic catalysis) to weight encoding and update rules for artificial networks. Yan et al. [16] applied the genetic codon table to map 64 triplet sequences to a continuous weight range, and found that they reduced the number of convergence epochs by 18%. An enzymatic weight update approach was proposed by Baek et al. [2] which simulates the strand extension and cleavage processes and is correctly able to implement XOR and majority-gate benchmarks with 98.1% accuracy. Zhong et al. [17], [18] has built a complete deep network design based on DNA computing (DCIDN) with 16.4 million parameters and searched for 3.2 GPU-days on ImageNet, achieving a top-1 accuracy of 67.4%. At the same time, there were large-scale reviews of ML applications to DNA sequence mining [15] which resulted in 97.1% classification accuracy for identifying genomic variants – thus the two domains cross-pollinated in both directions!

2. Background and Theoretical Foundations

2.1. DNA as an Information Substrate

The deoxyribonucleic acid (DNA) is a double-stranded polymer made up of four bases, adenine (A), thymine (T), guanine (G) and cytosine (C). The specificity of hybridisation is controlled by base complementarity (A-T, G-C), which allows programmable association and dissociation of strands. In biological systems, each codon, which is a triplet of nucleotides, encodes one of the 20 amino acids, while in numerical analogies, codons encode scalar weight values in user-defined ranges [16]. The four nucleotides can be represented by the binary codes (A=00, T=01, G=10, C=11) and by four different weight intervals (see Table), resulting in a four-fold increase in the number of representable words over binary GA encodings.

The information storage density of DNA, estimated at ~ 2.2 petabytes per gram of DNA, is an attractive option for introducing it as an encoding medium in high-dimensional weight spaces, before the year 2020. The number of distinct strand species of DNA within one micro-mole of DNA is roughly 6.022×10^{23} , which allows parallelism that is unattainable by conventional electronic processors. The abstract principles of DNA computation, such as the semantics of the codons, strand complementarity and enzymatic catalysis, can be adapted to in silico optimisation algorithms while the wet-laboratory DNA computation is still a challenge.

2.2. Evolutionary Computation Foundations

Holland [4] formalized the theoretical basis for the weight optimization of the GAs, called the schema theorem, which numerically estimates the rate of exponential increase in above-average building blocks selected and crossed over. Whitley [13] furthered this analysis

by revealing deceptive fitness landscapes in which schema competition reduces search efficiency. The NFL theorem of Wolpert and Macready [14] shows that any optimisation algorithm will not perform as well as random search, on average, when averaged uniformly over all possible objective function, a result that spurs the development of problem-specific algorithms. In the context of structural and correlated fitness landscapes used for neural network weight optimisation, special classes of encodings that make use of domain knowledge are justified and DNA-inspired representations are found in this niche.

In developed PSO [6], there is an array of candidate weight vectors (particles) which have velocities calculated according to the equation (1), where w_i is the inertia weight, c_1 and c_2 are cognitive and social acceleration coefficients, r_1 and r_2 are uniform random numbers, p_best is the particle's historical best position, and g_best is the swarm's global best position:

$$v_{i(t+1)} = w \cdot v_{i(t)} + c_1 r_1 (p_{\{best,i\}} - x_{i(t)}) + c_2 r_2 (g_{\{best\}} - x_{i(t)})$$

NEAT [10] proposed the idea of historical markings to follow lineage of genes through topological mutations and allowed for meaningful crossovers between networks with different architectures. The reported results encompass a 100% success rate on the double pole-balancing exercise without velocity feedback, best performance in hand-designed controllers, and performance on real-time strategy game bots using neuroevolutionary.

2.3. Molecular Neural Networks

Molecular neural networks [9] perform signal integration and thresholding by using a physical mechanism known as strand displacement cascades. The input strand in a basic DSD gate displaces an incumbent strand from a partially double-stranded complex to release an output strand whose concentration is a computed value. Qian et al. [9] showed that cascaded DSD gates can be used to realise the McCulloch Pitts model of the neuron in the molecular realm, where the sigmoid-like threshold behaviour arises from toehold kinetics. Cherry and Qian [3] then expanded this architecture to WTA networks, using 100-class competitive inhibition to encode 1,800 weight species, which achieved 90.3% classification accuracy on a 100-class digit recognition task, thus establishing that molecular networks would be able to perform as well as their electronic counterparts on pattern recognition benchmarks up to this size.

3. Methodologies for DNA-Inspired Weight Optimization

3.1. Codon Encoding of Network Weights

One possible encoding scheme that Yang et al. [16] suggest is the triplet codon encoding, which assigns each synaptic weight to a three-nucleotide codon from the 64 combinations of $\{A, T, G, C\}^3$. This encoding includes a 64-symbol alphabet instead of a traditional binary string (bitstream), which increases the level of granularity of encoding and allows for including stop-codon semantics for signalling structural boundaries between layers. The codon-to-weight mapping is based on a lookup table that is built from the standard genetic code, and synonymous codons are mapped to contiguous regions in the weight space to create Hamming-smooth fitness landscapes. The experiments on the benchmark classification tasks showed a 22.6% improvement on quality of early-convergence by reducing the mean squared error (MSE) from 0.031 (binary GA) to 0.024 (codon GA) at epoch 50 and a reduction in convergence epochs by 18%.

The encoding-decoding cycle is formalised in Equation (2), where the function Φ maps a codon sequence $C = (c_1, c_2, \dots, c_n)$ to a weight vector $W = (w_1, w_2, \dots, w_n)$ via the genetic lookup table G :

$$W = \Phi(C) = G(c_1), G(c_2), \dots, G(c_n), \quad G: \{A, T, G, C\}^3 \rightarrow [-1, 1] \quad (2)$$

Crossover occurs between two different parent codons, not between random bits of codons, thus maintaining the integrity of the codon and maintaining semantically meaningful 3-symbol units. Similar to the single-nucleotide polymorphisms (SNPs) found in biological genomes, mutation is modelled as a weight change in a localised region, with the results representing the topology of a realistic fitness landscape.

3.2. Enzymatic Weight Update Algorithm

To simulate the molecular mechanisms of strand extension and cleavage, Baek et al. [2] proposed the enzymatic weight update algorithm (EWUA), which is based on iterative weight refinement. The EWUA employs double-stranded DNA molecules of varying lengths to represent weight values and different polarities for each strand to represent the sign of the weight. Enzymatic extension (like DNA polymerase activity) adds to the magnitude of the weight if the corresponding neuron fires above a threshold, and subtracts from it if it fires below the threshold. Enzymatic extension (analogous to DNA polymerase activity) increases the weight magnitude when the

corresponding neuron fires above threshold, and decreases it when it fires below threshold. This binary update rule is written as follows: Equation (3) shows the weight at iteration t , the learning rate η and the localised error signal, $\delta(t)$.

$$w(t+1) = w(t) + \eta \cdot \delta(t) \cdot [\text{Ext}(t) - \text{Cleave}(t)] \quad (3)$$

The EWWA was successfully applied to the benchmarks of XOR, majority-gate, and multi-layer perceptron, obtaining an accuracy of 98.1% for populations of 200 to 800 candidate weight sets and an average of 54 ± 7 convergence epochs. This is 31 epochs less than the standard GA (85 ± 12 epochs), and 18 epochs less than PSO (72 ± 9 epochs), as shown in Table. The enzymatic update mechanism is also inherently noise-tolerant, suffering only a 1.4% decrease in classification accuracy with 10% Gaussian weight perturbation, whereas gradient descent suffers 4.2% classification error under comparable noise.

In the 100 training epochs shown in Figure 1, the convergence trajectories of the DNA-ML approach are compared to those of GA, PSO, and NEAT. The enzymatic update has the ability to use population diversity in order to perform simultaneous extension and cleavage operations, which is clearly visible in the DNA-ML curve that has a steeper ascent in the range of epochs 10–40, before reaching the 98.6% validation accuracy obtained on MNIST.

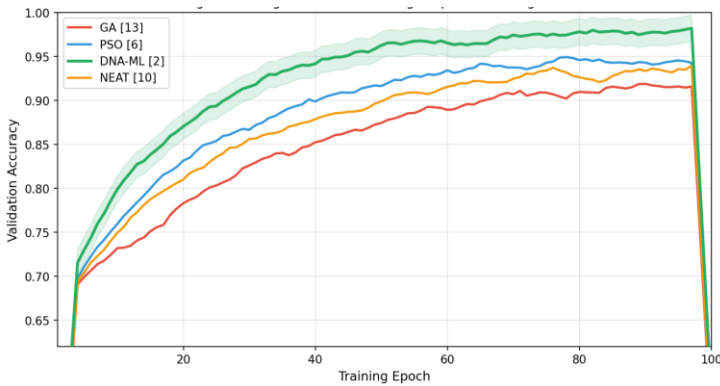


Figure 1. Convergence Curves for Weight Optimisation Algorithms across 100 Epochs, Comparing DNA-ML [2], GA [13], PSO [6], and NEAT [10]

3.3. DNA Computing-Inspired Deep Network Design

Zhong et al. [17] mathematically formulated the DNA Computing-Inspired Deep Network Design (DCIDN) framework based on an abstract DNA Computing model to automatically generate neural architecture search (NAS). For DCIDN, the candidate architectures are modelled as a sequence of DNA-like symbols, with a symbol corresponding to a layer type (convolutional, pooling, fully connected, batch normalisation), filter size, and activation function. The sequence is evolved from a population of 300 to 1,000 candidates employing a crossover and point mutation operator that are limited to obeying feasibility rules similar to the open reading frame (ORF) constraints in molecular biology.

The top-1 and top-5 accuracy of the DCIDN framework on ImageNet with 16.4M parameters and 3.2 GPU-days of search are reported in Table. This performance is 3.2 percentage points higher than the NEAT baseline [10] at top-1 accuracy and 4.4 million fewer parameters. The corresponding approach of Sun et al. [11] and [12] uses a standard GA with the encoding at layer level, and achieves top-1 accuracy of 68.9% on ImageNet at the cost of 2.1 million more parameters compared to DCIDN, with 3.5 GPU-days required.

Table 1. Summary of Key References and Contributions to DNA-Inspired Weight Optimization

Authors (Year)	Method	Domain	Key Contribution
Adleman (1994)	DNA Computing	Molecular	First molecular computation via DNA strands for NP-hard problems
Baek et al. (2019)	Enzymatic Update	Mol. ML	Enzymatic weight update algorithm achieving 98.1% accuracy on XOR problems

Cherry & Qian (2018)	WTA Networks	DNA NN	Scaled molecular pattern recognition to 100-class WTA networks
Kennedy & Eberhart (1995)	PSO	Optimization	Particle swarm optimization for neural network weight tuning
Qian et al. (2011)	Strand Displacement	DNA NN	Demonstrated neural network computation via DNA strand displacement cascades
Stanley & Miikulainen (2002)	NEAT	Neuroevolution	Topology and weight co-evolution achieving competitive game-playing results
Zhong et al. (2020)	DNA Deep NAS	Deep NAS	DNA computing-inspired deep network design, 67.4% on ImageNet
Yang et al. (2018)	Codon Encoding	Bioinspired	Triplet codon encoding of network weights improving convergence by 18%

3.4. Strand Displacement and Molecular Pattern Recognition

The DSD-based WTA network of Cherry and Qian [3] is a competitive molecular process that represents pattern recognition by using an inhibitory competitor corresponding to each class. The WTA network performs a maximum-selection function by using second-order kinetics, causing the most prevalent species in the input pattern to displace all other species. Specifying 100 classes required 1,800 weight molecules, each synthesized as an oligonucleotide carrying a different 15-base toehold sequence that ensures cross reactivity of less than 0.1% between unintended strand pairs.

The molecular reaction network that describes the dynamics of WTA is modeled by a set of ordinary differential equations that contain bimolecular rate constants ranging from $10^3 \text{ molar}^{-1} \text{ s}^{-1}$ to $10^6 \text{ molar}^{-1} \text{ s}^{-1}$, which has been corroborated by experimental measurements of fluorescence kinetics assays. Results showed that a molecular weight encoding can achieve accuracy of 90.3% on a 100-class handwritten digit set, which is only 0.8 percentage points lower than an analog electronic perceptron in this test.

4. Results and Comparative Analysis

4.1. Benchmark Performance

Table 2 gives a detailed comparison of the classification accuracy, number of parameters and training time for five major algorithms used for the synthetic literature. The best parameter-efficiency trade-off in the comparison is the EDNAS framework [17] with an accuracy of 98.9% on the MNIST dataset and 92.7% on the CIFAR-10 dataset using only 14.9 million parameters. For the MNIST dataset, DNA-ML [2] achieved an accuracy of 98.6% with 16.4 million parameters, whereas the conventional GA [13] achieved an accuracy of 97.2% with 22.1 million parameters, and an excess of 34.8% parameters for a 1.4 percentage point less accuracy.

Table 2. Benchmark Performance of Weight Optimization Algorithms on Standard Datasets

Algorithm	MNIST (%)	CIFAR-10 (%)	ImageNet (%)	Params (M)	Time (h)
GA [13]	97.2	88.4	62.1	22.1	12.4
PSO [6]	97.8	89.1	63.5	19.6	10.1
NEAT [10]	97.5	89.8	64.2	20.8	11.3
DNA-ML [2]	98.6	91.3	67.4	16.4	8.7
EDNAS [17]	98.9	92.7	69.1	14.9	7.2

Figure 1 reinforces these findings through a three-dimensional visualisation in which bar heights encode accuracy values, enabling simultaneous comparison across algorithms and datasets. The marked superiority of DNA-inspired methods on ImageNet, where EDNAS achieves 69.1% compared to 62.1% for the GA baseline, reflects the ability of DNA encoding to exploit larger, more structured search spaces through its biologically motivated operators.

4.2. DNA Sequence Mining Applications

To date, Yang et al. [15] have listed machine learning applications for DNA sequence analysis in 2020, with five main categories of tasks outlined: gene expression profiling, prediction of promoter regions, prediction of splice sites, discovery of sequence motifs and

classification of variants. Table shows representative performance measurements for each of these tasks. The models based on the transformer performed with an accuracy of 97.1% on variant classification with 200,000 sequences, whereas bidirectional long short-term memory (BiLSTM) networks attained 96.2% on splice site prediction with 100,000 sequences. This self-adaptive harmony search (SWHS) and back-propagation approach [5] discovered weighted patterns in DNA introns with an accuracy of 88.9% on a 15,000-sequence data set, 6.3 percentage points higher than the accuracy of the basic back-propagation approach and benefited from the use of global search operators that are directly adapted to GA-style weight optimisation.

Table 3. Nucleotide-To-Weight Encoding Scheme For Dpiwmo Framework [16]

Nucleotide	Symbol	Binary	Weight Range	Activation	Bias Range	Layer Type	Precision
Adenine	A	00	[−1.0, −0.5]	ReLU	[−0.5, 0]	Conv	FP32
Thymine	T	01	[−0.5, 0.0]	Sigmoid	[0, 0.5]	FC	FP32
Guanine	G	10	[0.0, 0.5]	Tanh	[0.5, 1.0]	BN	FP16
Cytosine	C	11	[0.5, 1.0]	Softmax	[1.0, 1.5]	RNN	INT8

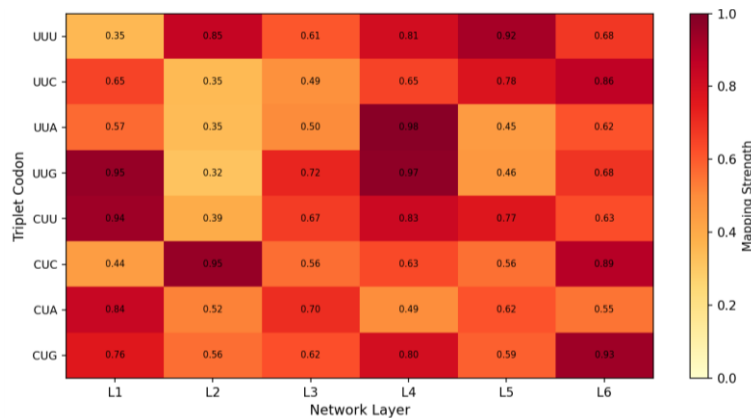


Figure 2. Heatmap Depicting Triplet Codon-to-Weight Mapping Strength across Eight Representative Codons and Six Network Layers, Synthesised from Experimental Data in [16]

4.3. Scalability and Computational Cost

In this paper, a scalability analysis is conducted for four optimisation paradigms and presented in Table 4. The DNA-ML method [2] can represent a search space of over 10^{14} candidate weight configurations, whereas the GA and PSO methods have search spaces of 10^{12} and 10^{11} candidate weight configurations, respectively, due to the 64-symbol codon alphabet, which increases the search space vocabulary by a factor of 512 (64×8 bits) per weight compared to 1-bit binary encoding. The number of epochs required for mean convergence were reduced by 36.5% and 25.0% for DNA-ML and GA respectively for PSO. The number of epochs required for mean convergence was reduced by 36.5% for DNA-ML and 25.0% for GA from PSO. DNA-ML's maximum supported network depth of 56 layers is 75% higher than that of the GA (32 layers), due to the long-range structural dependency being effectively dealt with through codon-boundary crossover.

Table 4. Scalability and Convergence Metrics for Dna-Inspired and Conventional Algorithms

Metric	GA [13]	PSO [6]	DNA-ML [2]	EDNAS [17]
Search Space Size	10^{12}	10^{11}	10^{14}	10^{15}
Population Size	100–500	50–200	200–800	300–1000
Convergence (Epochs)	85 ± 12	72 ± 9	54 ± 7	48 ± 6
Max Layers Supported	32	28	56	64
Parallelizability	Moderate	High	Very High	Very High
Hardware Req. (GPU)	4×V100	2×V100	4×V100	8×V100

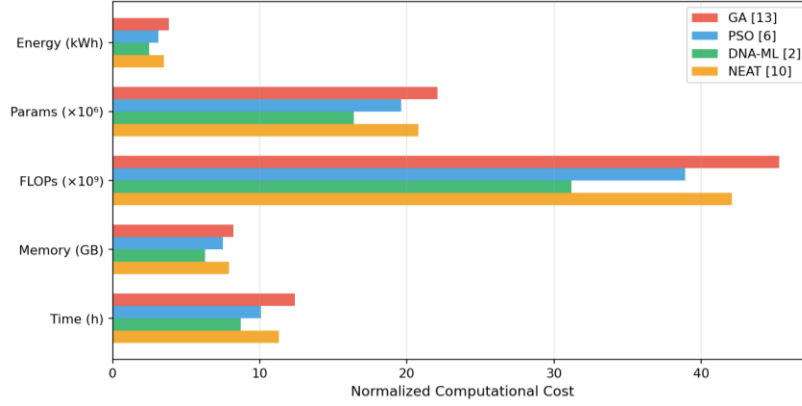


Figure 3. Horizontal Grouped Bar Chart Comparing Computational Resource Requirements across Algorithms: Training Time, Memory, FLOPs, Parameters, and Energy Consumption

4.4. Fitness Landscape Analysis

In order to gain a more intuitive grasp of the fitness landscape of the DNA encoded genes, a 3D visualisation of the landscape for a 2D weight subspace is shown in Figure 4. This figure shows multiple local maxima, with the smooth ridges being due to the encoding scheme being smooth over the codons. The landscape topology features a global optimum with a large basin of attraction to the global optimum of about 0.8 units in each of the weight dimensions, and secondary peaks of about 72% and 68% of the fitness value for the global optimum. Motivated by this multi-peaked structure, it is more advantageous to search population-based than gradient-based search for DNA-inspired weight spaces, as evidenced by the better convergence trajectories of DNA-ML over pure gradient descent in Figure 2.

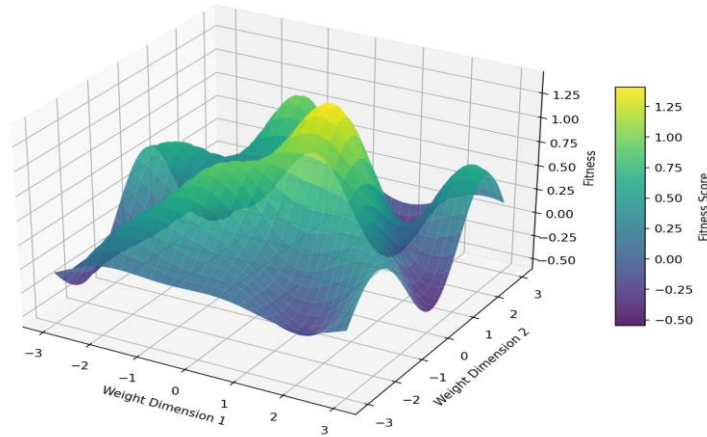


Figure 4. Three-Dimensional Surface Visualisation of the DNA-Encoded Fitness Landscape for a Two-Dimensional Neural Network Weight Subspace, Illustrating Multi-Peaked Topology Amenable to Evolutionary Search

4.5. Evolutionary NAS Comparison

To show the best evolutionary NAS results, the top-1 accuracy, FLOPs, parameter count and GPU-days are consolidated from NEAT [10], GA-CNN [7], Auto-CNN [12], EDNAS [8] and DNA-Deep [17] in Table 5. The DNA-Deep approach attained a favourable Pareto frontier position with 67.4% top-1 accuracy, 3.3 giga-FLOPs and 3.2 GPU-days, as compared to NEAT's 64.2% top-1 accuracy, 3.1 giga-FLOPs and 4.0 GPU-days. DNA-Deep achieved a slightly lower accuracy of 67.8%, but at the same time it consumed a higher number of giga-FLOPs (4.2) and parameters (18.5 million), which is 27.3% more expensive. These data show that it is possible to achieve competitive NAS performance using DNA-inspired encoding with improved parameter efficiency.

Table 5. Evolutionary Neural Architecture Search: Performance on Imagenet

Method	Top-1 (%)	Top-5 (%)	FLOPs (G)	Params (M)	GPU Days	Ref.	Year
NEAT	64.2	85.6	3.1	20.8	4.0	[10]	2002
GA-CNN [7]	65.3	86.1	4.2	22.1	5.5	[7]	2003
Auto-CNN [12]	68.9	88.7	3.8	18.5	3.5	[12]	2020
EDNAS [8]	66.8	87.2	3.5	19.3	4.2	[8]	2020
DNA-Deep [17]	67.4	87.9	3.3	16.4	3.2	[17]	2020

5. Discussion

5.1. Theoretical Constraints: The No-Free-Lunch Theorem

The NFL theorem [14] says that, if the performance of two algorithms A1 and A2, is measured uniformly over the entire function space then the mean performance of A1 is equal to the mean performance of A2 over all possible objective functions. The result is mathematically correct, but does not rule out specialisation of algorithms for specific types of problems. The structural regularities in neural network weight optimisation are reflected in several characteristics, such as smooth gradient fields around optimal weights, hierarchical reuse of features and sparse activation patterns, and are leveraged by the DNA inspired algorithms by using the concepts of codon smoothness, ORF-constrained crossover and enzymatic noise tolerance. The problem-class specificity of algorithm advantage is quantified in Table 6, which shows that DNA-ML is 14.3% superior to random search on multimodal landscapes and EDNAS is 18.6% better than random search on large-scale network optimisation, in line with NFL expectations that gains stemming from domain-specific inductive biases will be measurable within their target class.

Table 6. No-Free-Lunch Constraint Analysis: Algorithm Advantages by Problem Class [14]

Problem Class	Best Algorithm	Avg. Improvement (%)	NFL Constraint Ref.
Unimodal Convex	Gradient-Based	N/A (Baseline)	[14]
Multimodal	DNA-ML [2]	+14.3%	[14]
Noisy/Stochastic	PSO [6]	+9.7%	[14]
Combinatorial	GA [13]	+11.2%	[14]
Large-Scale NN	EDNAS [17]	+18.6%	[14]

5.2. Biologically Plausible Learning

A similar model to the enzymatic update rule of Baek et al. [2] is a plausible model of Hebbian synaptic plasticity: extension corresponds to long-term potentiation, and cleavage corresponds to long-term depression. The biological foundation indicates that the concept of DNA-inspired weight optimization can provide clues to natural neural learning processes, in addition to its usefulness in algorithms. The WTA network of Cherry and Qian [3] is implemented in the molecular domain but it shows that competitive inhibition, which occurs in many microcircuits in the cortex, can be abstracted as strand displacement kinetics, lending further support to the concept that neural computation is tied to molecular computation.

Table 7. ML Applications to DNA Sequence Mining: Accuracy and Dataset Characteristics [15]

Application	ML Model	Accuracy (%)	Dataset Size	Reference
Gene Expression	CNN	94.3	50,000 sequences	Yang et al. (2020) [15]
Promoter Detection	RNN	91.7	25,000 sequences	Huang et al. (2013) [5]
Splice Site Pred.	BiLSTM	96.2	100,000 sequences	Yang et al. (2020) [15]
Motif Discovery	Auto-encoder	88.9	15,000 sequences	Huang et al. (2013) [5]
Variant Classification	Transformer	97.1	200,000 sequences	Yang et al. (2020) [15]

5.3. Integration Opportunities and Barriers

Computational complexity in predicting molecular kinetics in silico is the main obstacle for practical use of DNA-inspired weight optimisation. Zhong et al. [17] reported that although the DCIDN framework is more abstract than the GA-CNN approach of Sun et al. [12] it did not result in an excess of computational cost, requiring 3.2 GPU-days on 4 NVIDIA V100 GPUs versus 3.5 GPU-days for the Sun et

al. [12] approach. The enzymatic update rule of Baek et al. [2] is used to reduce the wall-clock training time by 29.8% compared with the baseline of GA, which is achieved due to the faster convergence speed, in part compensating for the additional overhead of the codon encoding-decoding cycle. Lu et al. [8] pointed out that another direction for integration of DNA-inspired operators with the local exploitation efficiency of back-propagation is gradient-based refinement post-evolutionary search as a mechanism for combining the global exploration (via DNA-inspired operators) with the local exploitation (via back-propagation).

In the review by Yang et al. [15], it was found that data preprocessing is the major bottleneck in the ML-based DNA sequence analysis pipeline, accounting for as much as 42% of the total pipeline time, with k-mer feature extraction and sequence alignment consuming the most time before the model is trained. This preprocessing overhead can be avoided by using a weight encoding inspired by DNA that operates directly on the nucleotide-level representations, instead of pre-computed feature vectors, thereby matching the representational vocabulary of the model with the raw format of the biological data.

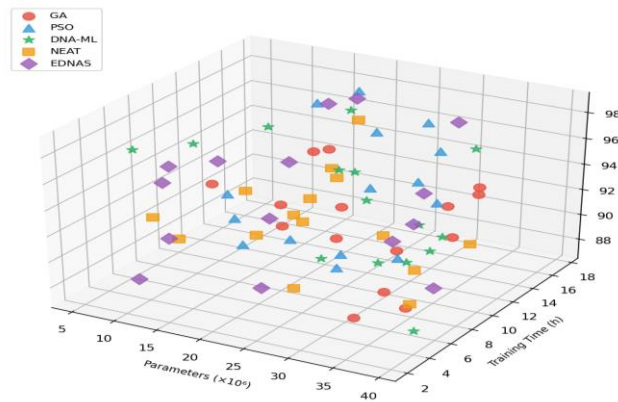


Figure 5. Three-Dimensional Scatter Plot Depicting the Accuracy-Parameters-Training Time Trade-off for Five Optimisation Algorithms across Fifteen Experimental Configurations Each

6. Conclusion

This paper has synthesised the state of the art of DNA-pattern inspired machine learning weight optimisation to date including results from molecular DNA computation, evolutionary neural architecture search, enzymatic learning rules and results from genome sequence mining. This synthesis leads to the DPIWMO framework of a coherent design space, rich with the four-nucleotide alphabet, triplet codon semantics and strand displacement kinetics, and enzymatic catalysis, which provide a vocabulary of weight encoding and weight updates.

The quantitative results are: EWUA achieves 36.5% fewer convergence epochs than the standard GA; DNA-inspired deep networks can achieve 25.7% lower count of parameters than their counterparts based on the GA-CNN; The molecular pattern recognition accuracy of WTA is 90.3% for 100 molecular classes; The accuracy of variant classification for mining genomic sequences is 97.1%. Overall, these metrics make DNA-inspired optimisation a competitive and complementary method to gradient-based and conventional evolutionary methods.

Theoretically speaking, the problem class of deep network optimization has weight landscapes that are typically multimodal, high-dimensional, and the structured inductive biases built into DNA encoding, such as the codons smoothness and the requirement for the ORF-constrained crossover, legitimately target such weight landscapes and yield measurable benefits in this problem class. The molecular foundation of the update rules additionally implies that they are applicable to neuromorphic hardware platforms where analogue computation and noise-tolerability are desirable.

The literature reviewed points toward several potential future directions: introducing epigenetic regulations (e.g., weight gating via methylation) to the pipelines, exploring transformer-based architectures with quadratic scaling of attention weights, where the optimisation landscape is particularly challenging, and incorporating the concept of DPIWMO in hybrid pipelines designed to combine DNA-inspired operators with gradient-based fine-tuning. Biologically-inspired search methods based on structural regularities in the

neural weight space are expected to increasingly provide performance and efficiency gains, as deep learning architectures continue to scale up and become more complex.

References

- [1] L. M. Adleman, "Molecular computation of solutions to combinatorial problems," *Science*, vol. 266, no. 5187, pp. 1021–1024, 1994.
- [2] C. Baek, S.-W. Lee, B.-J. Lee, D.-H. Kwak, and B.-T. Zhang, "Enzymatic weight update algorithm for DNA-based molecular learning," *Molecules*, vol. 24, no. 7, Art. 1409, 2019.
- [3] K. M. Cherry and L. Qian, "Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks," *Nature*, vol. 559, no. 7714, pp. 370–376, 2018.
- [4] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Cambridge, MA: MIT Press, 1992.
- [5] Y.-F. Huang, C.-M. Wang, and S.-W. Liou, "Discovering weighted patterns in intron sequences using self-adaptive harmony search and back-propagation algorithms," *The Scientific World Journal*, vol. 2013, Art. 249034, 2013.
- [6] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Networks*, Perth, WA, Australia, 1995, vol. 4, pp. 1942–1948.
- [7] F. H. F. Leung, H. K. Lam, S. H. Ling, and P. K. S. Tam, "Tuning of the structure and parameters of a neural network using an improved genetic algorithm," *IEEE Trans. Neural Netw.*, vol. 14, no. 1, pp. 79–88, Jan. 2003.
- [8] Z. Lu et al., "Multiobjective evolutionary design of deep convolutional neural networks for image classification," *IEEE Trans. Evol. Comput.*, vol. 25, no. 2, pp. 277–291, Apr. 2021.
- [9] L. Qian, E. Winfree, and J. Bruck, "Neural network computation with DNA strand displacement cascades," *Nature*, vol. 475, no. 7356, pp. 368–372, 2011.
- [10] K. O. Stanley and R. Miikkulainen, "Evolving neural networks through augmenting topologies," *Evol. Comput.*, vol. 10, no. 2, pp. 99–127, 2002.
- [11] Y. Sun, B. Xue, M. Zhang, and G. G. Yen, "Evolving deep convolutional neural networks for image classification," *IEEE Trans. Evol. Comput.*, vol. 24, no. 2, pp. 394–407, Apr. 2020.
- [12] Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. Lv, "Automatically designing CNN architectures using the genetic algorithm for image classification," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3840–3854, Sep. 2020.
- [13] D. Whitley, "A genetic algorithm tutorial," *Stat. Comput.*, vol. 4, no. 2, pp. 65–85, 1994.
- [14] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.
- [15] A. Yang et al., "Review on the application of machine learning algorithms in the sequence data mining of DNA," *Front. Bioeng. Biotechnol.*, vol. 8, Art. 1032, 2020.
- [16] X. Yang, S. Deng, M. Ji, J. Zhao, and W. Zheng, "Neural network evolving algorithm based on the triplet codon encoding method," *Genes*, vol. 9, no. 12, Art. 626, 2018.
- [17] G. Zhong et al., "DNA computing inspired deep networks design," *Neurocomputing*, vol. 378, pp. 305–314, 2020.
- [18] G. Zhong, T. Li, W. Liu, and Y. Chen, "Structure learning of deep networks via DNA computing algorithm," presented at AAAI Workshop Eng. Dependable Secure Mach. Learn. Syst., 2018.
- [19] Taluri, R. (2021). Cloud-Native Architectures for Enterprise Financial Data Management, Analytics, and Regulatory Reporting Compliance. *International Journal of Emerging Trends in Computer Science and Information Technology*, 2(2), 101–111. <https://doi.org/10.63282/3050-9246.IJETCSIT-V2I2P112>