Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Neuro-Hormonal Networks: A Bio-Inspired Architecture forEnhanced Deep Learning Performance

Furqan Yaqub Khan

2023phacse003@nitsir.ac.in

National Institute of Technology Srinagar

Tawseef Ayoub Shaikh

National Institute of Technology Srinagar

Research Article

Keywords: Neuro-hormonal Networks, Bio-inspired Computing, Neuro-modulation, De lay Differential Equations, Adaptive Neural Architectures, Hormonal Computing

Posted Date: March 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9029917/v1

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

Additional Declarations: No competing interests reported.

Neuro-Hormonal Networks: A Bio-Inspired Architecture for Enhanced Deep Learning Performance

Furqan Yaqub Khan* furqan_2023phacse003@nitsri.ac.in
Department of Computer Science and Engineering
National Institute of Technology
Srinagar, J&K 190006, India

Tawseef Ayoub Shaikh* tawseef.shaikh@nitsri.ac.in
Department of Computer Science and Engineering
National Institute of Technology
Srinagar, J&K 190006, India

Editor:

Abstract

In this work, neuro-hormonal networks (NHNs) are presented which are a novel bio-inspired deep learning architecture that integrates dynamic neuro-modulation mechanisms to achieve superior performance in deep learning tasks. Our approach addresses fundamental limitations in conventional neural networks and how they are enhanced by incorporating dopamine and cortisol-like hormonal pathways that adaptively modulate synaptic weights during training, mirroring biological neuro-endocrine systems. The mathematical framework is grounded in delay differential equations, control theory, and weighted logic networks, implementing three distinct modulation types: accelerators for enhanced plasticity, suppressors for homeostatic regulation, and stabilizers for output regularization. The multi-hormonal modulation component dynamically adjusts effective weights. Thus, enabling adaptive learning that responds to both current performance and historical training dynamics. Comprehensive evaluation across four benchmark datasets demonstrates consistent and statistically significant improvements: CIFAR-10 (+5.36pp), CIFAR-100 (+6.57pp), Fashion-MNIST (+0.67pp), and SVHN (+0.09pp). McNemar's test validation confirms statistical significance (p ¡ 0.001) for three of four datasets, with particularly strong benefits observed for complex, multi-class classification tasks. Advanced performance metrics include ROC AUC (≥ 0.99) and PR AUC (≥ 0.95) demonstrating superior discrimination capability and robustness across evaluation scenarios. The architecture exhibits enhanced convergence stability, improved generalization capability, and interpretable adaptation patterns that correlate with training dynamics. While incurring an average of 2.96× computational overhead per epoch, the biological plausibility and consistent performance gains validate the practical effectiveness of neuro-endocrine principles in artificial intelligence. Our work establishes a rigorous mathematical foundation for hormonal computing and demonstrates how multi-timescale biological dynamics can be successfully translated into improved deep learning architectures. **To the best of our knowledge, the NHN is the first novel work to establish and implement the neuro-hormonal modulation in deep learning architectures with rigorous mathematical background, advanced implementations and novel results.**

Keywords: Neuro-hormonal Networks, Bio-inspired Computing, Neuro-modulation, Delay Differential Equations, Adaptive Neural Architectures, Hormonal Computing.

1 Introduction

The field of deep learning (DL) has witnessed remarkable progress in recent years, yet conventional neural network architectures often struggle with adaptability, biological plausibility, and inefficient learning dynamics (Atzmueller and Lemmerich, 2019; Calisi and Saldanha, 2015). Biological neural systems employ sophisticated neuro-modulatory mechanisms that operate across multiple timescales, where fast neural signaling (milliseconds) is complemented by slower hormonal messaging (minutes to hours) (Brinkschulte, 2004; Wilson et al., 2018). This multi-timescale approach enables biological systems to exhibit remarkable adaptability, learning efficiency, and robustness to environmental perturbations (Encyclopædia_Britannica, 1998). The endocrine system, particularly through hormones such as dopamine and cortisol, provides a distributed control mechanism that modulates synaptic plasticity, learning rates, and behavioral responses across different contexts (Guerrero-Criollo et al., 2023). Recent advances in hormonal computing have demonstrated the potential for incorporating bio-inspired modulation mechanisms into artificial intelligence systems (Holcman and Triller, 2006). Mathematical foundations for such systems draw heavily from delay differential equations (DDEs), which capture the temporal dynamics inherent in hormonal feedback loops (Segel and Slemrod, 2013). The integration of hormonal modulation into neural networks requires careful consideration of weighted logic networks, which can be formalized using directed acyclic graphs (DAGs) with edge weights representing hormonal influence strength (Walker et al., 2010). Building upon control theory principles, neuro-endocrine systems can be modeled as proportional-integral-derivative (PID) controllers with hormonal modulation (Farkas et al., 2005). Recent developments in neural delay differential equations (NDDEs) have shown that incorporating temporal delays can significantly enhance the representational capacity of continuous-depth neural networks (Stagkourakis et al., 2020). These findings support the biological plausibility of our approach, as hormonal systems naturally exhibit delayed responses and memory effects that contribute to learning and adaptation (Dietterich, 1998). Current limitations in conventional deep learning architectures include insufficient biological plausibility, limited adaptability to changing environments, and sub-optimal learning efficiency in resource-constrained scenarios (Netzer et al., 2011). Bio-inspired approaches, particularly those incorporating neuro-modulation principles, offer promising solutions to these challenges by enabling more flexible, efficient, and robust learning mechanisms (Farkas et al., 2005). The mathematical framework for ultradian rhythms in hormonal systems can be modeled using coupled oscillators. Cross-modal integration in attention mechanisms can be enhanced through hormonal weighting.

1.1 Contributions and Objectives

The present work introduces a novel neuro-hormonal network (NHN) architecture that embeds biologically-inspired hormonal signals into deep learning. In our model, multiple hormone-like modulators (e.g. analogues of dopamine, cortisol, serotonin) dynamically influence synaptic weight updates, yielding context-dependent plasticity. Mathematically, we formulate hormone interactions using delay differential equations and control-theoretic constructs (e.g. hormone-modulated PID control) to capture the temporal feedback characteristic of endocrine systems. Through this framework, we implement a multi-hormonal

modulation mechanism that adjusts learning rates and neuron activations under different internal "hormonal" states. Experimentally, our NHN demonstrates significant improvements in classification accuracy on standard vision benchmarks (CIFAR-10, CIFAR-100, Fashion-MNIST, SVHN) compared to conventional models. These gains are validated by appropriate statistical tests (e.g. McNemar's test) to ensure robustness of the results. Importantly, our approach maintains neuro-biological plausibility: hormone-modulated weight updates produce slow or diffused changes reminiscent of real neuro-endocrine adaptation, endowing the network with enhanced dynamic adaptability and stability.

The scientific and empirical objectives of this research are fourfold.

- First, we aim to advance biologically-plausible neural architectures by explicitly incorporating endocrine-like control signals into network learning.
- Second, we seek to enhance adaptability, learning efficiency, and robustness of deep models through hormone-like feedback loops that guide synaptic plasticity in a context-sensitive manner.
- Third, we investigate the role of temporal delays and feedback in neural control systems, using delay differential equations and control theory to analyze how time-lagged hormonal signals can improve stability and response.
- Finally, we strive to provide a unified mathematical and computational foundation for neuro-endocrine-inspired learning models, bridging differential-equation formulations with modern deep learning. By doing so, this work lays the groundwork for future exploration of neuro-hormonal mechanisms in AI and contributes a principled framework unifying hormonal control theory with neural network design.

So in this work, we present a comprehensive evaluation of neuro-hormonal networks across four benchmark datasets: CIFAR-10, CIFAR-100, Fashion-MNIST, and SVHN, demonstrating consistent improvements in overall convergence speed, final accuracy, and robustness compared to architecturally equivalent standard convolutional neural networks. Our contributions include the development of a novel bio-inspired architecture, mathematical formalization of hormonal modulation mechanisms, comprehensive empirical evaluation with statistical significance testing, and demonstration of improved learning dynamics and generalization capabilities.

1.2 Layout

The rest of the paper is divided into seven sections starting with Section 2 till Section 8. Sections 2 (*Literature Review*) starts with bio-inspired neural networks followed by neuromodulation and delay differential equations in neural systems and then discusses feedback control mechanism and cross model attention mechanism with hormone weighing and ends with control theory applications of artificial intelligence. Section 3 (*Mathematical Framework*) consists of foundations mathematical background for delay differential equations and weighted logic networks followed by proportional integral derivative control with hormonal modulation then by multi-timescale dynamics and ultradian rhythms and then by multi-hormonal modulation architecture and hormone based update dynamics and then finalizes

with training protocol and optimization criteria. Section 4 (*NHN Architecture*) deals with how the mathematical logic has been transformed into programming logic thus starting with multi-hormonal modulation mechanism then network architecture design then finally by training protocol. Section 5 (*Experimental Setup*) is with three subsections namely datasets and preprocessing, implementation details and evaluation metrics. Section 6 (*Result and Analysis*) provides with performance evaluation, statistical significance testing, computational overhead analysis and learning dynamics analysis. section 7 (*Discussion*) provides overall view of performance of proposed architecture and Section 8 (*Conclusion and Future Directions*) concludes the paper as well provides way for further research.

2 Literature Review

The bio-inspired neural network field has experienced a surge in the past several years that involves researchers mining biology for neural modulators to boost artificial intelligence systems. The theoretical motivations, the mathematical formalism, and the empirical insights which have paved the way towards the current state-of-the-art of neuro-hormonal computing architectures are then reviewed. The literature review is divided into five major sections, each addressing a distinct aspect of neuro-hormonal networks and their theoretical, mathematical, and computational foundations. The structure has been designed to progress from conceptual underpinnings to applied mechanisms and mathematical background with benchmarks in the field thereby ensuring clarity and systematic coverage of the field.

2.1 Bio-Inspired Neural Networks

The conceptual model of hormone computing is a departure from the usual neural network and it is bio-inspired, with multi timescale dynamics drawing on biologic knowledge offering wide guidance for the deployment of hormonal bio-inspired models in various artificial intelligence (AI) applications by putting forth four main directions: bio-inspired hormonal computer systems, bridges to living systems, programming languages, and hybrid human-machine systems (Vallverdu et al., 2023). The present investigation is a first step to demonstrate that the endocrine system represents a robust multi-dimensional information processing architecture shaping the neural dynamics and controlling parallel multitasks through distributed control (Lagani et al., 2023). Early work on hormonally modulated neural control established the basic principle that synapses have sensitivity terms that modulate their responsiveness to specific hormone concentrations (Brinkschulte, 2004). This seminal work proposed the notion of "Autologous hormone control of Pagurian behavior" in which feedback loops control the hormone balance, so that artificial glands are stimulated by sensor input from the environment to generate the hormone quantities directly related to the quality of the sensor input. The sensitivity-based synaptic connections can alter the dynamic behavior of the network (Brinkschulte, 2004). The integration of cellular neuro-modulation principles into deep neural networks addresses fundamental limitations in adaptation and continual learning. Guerrero-Criollo et al. introduce bio-inspired neural networks that incorporate cortical synaptic circuits, short-term memory circuits, winner-take-all competitive networks, and nonlinear oscillation circuits for autonomous navigation tasks. This approach demonstrates significant improvements in decision-making mechanisms and motor control modulation (Guerrero-Criollo et al., 2023). Contemporary research

extends neuro-modulation concepts to spiking neural networks, where heterogeneous neuron populations generate diverse dynamic modes enabling inherent flexibility. Mei et al. propose neuro-modulation-aware SNNs that integrate neuron-specific neuro-modulation with heterogeneous architectures, enabling context-aware signal adaptation (Mei et al., 2025).

2.2 Neuromodulation in Deep Learning

Advanced neuro-modulation research explores multi-neuro-modulatory interactions that enrich single-neuro-modulator-driven learning (Vallverdu et al., 2023). The mathematical framework incorporates multiple neuro-modulators including dopamine (DA), acetylcholine (ACh), serotonin (5-HT), and noradrenaline (NA), operating across multiple spatial and temporal scales (Lagani et al., 2023). This approach demonstrates a "many-to-one" neuro-modulator-to-task mapping that enhances robustness and reduces catastrophic forgetting. The mathematical modeling of synaptic weight modulation incorporates receptor trafficking dynamics and scaffolding protein interactions (Holcman and Triller, 2006). The structure provides capacity to change the weights dynamically over the hormone state variables. The time-dependent hormone rates decay with continuous updates are used. This component introduces the effects of regulatory signals or environmental influences on hormone production, effectively balancing decay with replenishment. Together, the exponential decay and input-driven replenishment provide a framework for modeling dynamic hormonal responses under both basal and stimulated conditions. They also form the mathematical foundation of biologically plausible learning rules (Vallverdu et al., 2023).

2.3 Delay Differential Equations in Neural Systems

Neural delay differential equations (NDDEs) are a breakthrough in continuous-depth neural networks with improved representational capacity over standard neural ordinary differential equations (NODEs) (Zhu et al., 2023). Zhu et al. prove that NDDEs have universal approximation capabilities. The adjoint sensitivity technique for gradient computation in NDDEs allows efficient training of these delay-based models (Zhu et al., 2021). The mathematical theory shows that differential equations with time delays are dynamical systems of infinite dimension, exhibiting more productive dynamics and larger nonlinear representation ability compared to traditional methods (Vinodbhai and Dubey, 2023).

2.3.1 Biological Inspiration from Multi-Timescale Dynamics

Biological neural systems operate across multiple temporal scales, with fast neural signaling occurring in milliseconds while hormonal messaging spans minutes to hours. Mochida et al. develop neural networks incorporating both fast and slow dynamics inspired by cortical timescale hierarchies, demonstrating that slow dynamics store input-output history and influence fast dynamics based on accumulated memory (Mochida et al., 2021). The multi-timescale systems employs coupled differential equations where slow variables act as bifurcation parameters for fast dynamics. This framework allows the modeling of systems in which interactions between fast and slow processes generate complex behaviors, including oscillations, bursts, and other non-Markovian sequences. By incorporating history-dependent state transitions, the fast-slow system captures temporal dependencies and memory effects that are often observed in biological, chemical, and neural networks,

enabling a more realistic representation of multiscale dynamical phenomena (Mochida et al., 2021).

2.3.2 Ultradian Rhythms and Oscillatory Networks

Recent research reveals the presence of ultradian calcium rhythms in hypothalamic regions, providing insights into the mathematical modeling of biological oscillatory systems (Honma et al., 2018). The paraventricular nucleus (PVN) and subparaventricular zone (SPZ) exhibit synchronous ultradian rhythms with periods ranging from 0.5 to 4.0 hours. Hill-type kinetics are commonly used to model ultradian rhythms and feedback-driven biological processes, where cooperative regulation and degradation jointly shape the temporal dynamics (Honma et al., 2018). These oscillatory dynamics provide the mathematical foundation for implementing hormone-based scheduling in artificial neural networks.

2.4 Feedback Control Mechanism and Cross Model Attention with Hormone Weighing

Adaptive hormone regulation employs gradient-based feedback control. The model provides a flexible framework to describe feedback-driven hormonal regulation, highlighting the capacity of biological systems to fine-tune responses based on error signals and context-dependent sensitivities (Bertram et al., 2015). The automatic adjustment of hormonal levels according to learning progress and error signals is made possible through this mechanism. Recent advances in attention mechanisms incorporate hormonal weighting to enhance cross-modal processing capabilities (Vallverdu et al., 2023). This method allows for the dynamic modulation of attention by hormonal state variables. Studies on cross-modal attention show that attention mechanisms can facilitate as well as constrain processing across sensory modalities. The addition of hormonal modulation furnishes a mathematical structure that allows the construction of biologically plausible attention control by minimizing a cost function corresponding to environmental demands.

2.5 Control Theory Applications in AI

Hormonal influence networks are formally described mathematically as directed acyclic graphs (DAGs) where the weights on the edges represent hormone-specific modulation strength (Vallverdu et al., 2023). This framework enables the implementation of neuro-endocrine-PID controllers that integrate control theory principles with biological hormone modulation (Davis and Goadrich, 2006). This combination of proportional, integral, derivative, and hormonal feedback allows the controller to respond dynamically to both immediate and delayed changes in the system, making it suitable for modeling biological regulatory processes such as hormonal control loops (Fitzgerald et al., 2020). This integration enables adaptive control mechanisms that respond to both error signals and hormonal state variables (Bertram et al., 2015). The application of hormone-inspired control extends to various engineering domains, including artificial pancreas systems where PID controllers incorporate insulin feedback mechanisms. These implementations demonstrate hormone-like feedback control that prevents over-delivery and enhances safety through physiologically inspired regulation (Doyle et al., 2015). Neural network control theory applications, particularly in architectures that unify recurrent networks, spiking neurons, and deep learn-

ing approaches, provide improved stability, better generalization capability, and reduced computational complexity compared to traditional architectures. This integration provides mathematical frameworks for modeling complex regulatory systems.

2.5.1 Performance Advances and Open Challenges

Empirical evaluation of neuro-hormonal networks demonstrates consistent performance improvements across multiple benchmark datasets. Lagani et al. provide comprehensive survey evidence showing that bio-inspired approaches, including neuro-modulation-based methods, achieve superior performance in various machine learning tasks (Lagani et al., 2023). The mathematical frameworks developed for hormonal computing enable these performance gains through enhanced adaptability and biological plausibility (Vallverdu et al., 2023). Recent real-time AI implementations in neuro-modulation achieve system response times under 3 milliseconds, demonstrating the practical feasibility of hormone-based control systems. These advances integrate mainstream AI algorithms with neural signal processing, enabling comprehensive capabilities including neural sensing, edge AI computing, and electrical stimulation. The literature reveals several promising research directions for neuro-hormonal computing. Phase-locked loop oscillatory neural networks offer novel approaches to frequency clustering and synchronization problems (Doyle et al., 2015). These systems demonstrate that symmetric synapse matrices enable convergence to stable phase relationships dependent on synaptic weights. Weighted Signal Temporal Logic Neural Networks (wSTL-NN) represent emerging neuro-symbolic frameworks that combine neural networks with temporal logics (Zhu et al., 2021). These approaches enable end-to-end differentiable learning of temporal formulas through back-propagation, providing more explainable features than classical methods. The integration of Hill function-based dynamic models with neural networks provides mathematical frameworks for modeling complex regulatory systems. These approaches enable parameter estimation for under-determined optimization problems through separation of threshold and co-operative parameters. Recent advances in mathematical modeling of hypothalamic-pituitary-adrenal axis dynamics incorporate hippocampal receptors, arginine vasopressin signaling pathways, circadian rhythm functions, and short negative feedback loops to enhance predictive accuracy. These developments demonstrate the importance of comprehensive biological modeling in understanding neuro-endocrine system dynamics.

3 Mathematical Framework

This section is further divided into 8 subsections as discussed below.

3.1 Delay Differential Equations Foundation

The temporal dynamics of hormonal systems are mathematically modeled using delay differential equations (DDEs) given in equation (1):

$$\frac{dX}{dt} = f(X(t), X(t-\tau)) \tag{1}$$

Where $\frac{dX}{dt}$ represents the rate of change of the state variable X with respect to time t. In the context of the hypothalamic–pituitary–adrenal (HPA) axis, $X(t)$ may correspond to the concentration of a hormone such as cortisol, adrenocorticotropic hormone (ACTH), or corticotropin-releasing hormone (CRH). The term $X(t)$ denotes the current state of the system at time t, reflecting the present hormone concentration in the bloodstream, while $X(t-\tau)$ represents the delayed state of the system, corresponding to hormone levels at an earlier time $t-\tau$, thereby capturing the biological reality that endocrine feedback is not instantaneous. The parameter τ serves as the feedback delay, accounting for the time required for hormone secretion, circulation, receptor binding, and subsequent feedback effects; for the HPA axis, this delay typically lies in the range of 40–60 minutes. Finally, the function $f(X(t), X(t-\tau))$ characterizes the regulatory mechanism, with its dependence on $X(t)$ capturing immediate regulation and its dependence on $X(t-\tau)$ modeling delayed negative or positive feedback, thus enabling a physiologically realistic description of hormonal dynamics.

3.2 Weighted Logic Networks

Direct acyclic graphs (DAGs) with weighted edges reflect hormonal influence networks given in equation (2):

$$H_i(t+\Delta t) = \sigma\left(\sum_j W_{ij} \times H_j(t) + I_{ext}\right) \tag{2}$$

where $H_i(t+\Delta t)$ represents the updated state of the i-th node or unit at time $t+\Delta t$, and $H_j(t)$ denotes the current state of the j-th node. The term W_{ij} corresponds to the weight of the connection from node j to node i, capturing the strength and influence of the interaction between these nodes. The summation over j accumulates the weighted contributions from all connected nodes, while I_{ext} represents any external input applied to the system, such as a stimulus or control signal. The function $\sigma(\cdot)$ is a nonlinear activation function that maps the total input to a suitable output range, often introducing saturation or threshold effects in the system. This formulation allows the dynamics of each node to depend on both the network interactions and any external influences, enabling the modeling of complex feedback, cooperative, and regulatory behaviors in networks such as neural circuits or signaling pathways. By updating the states iteratively over discrete time steps Δt.

3.3 Proportional Integral Derivative (PID) Control with Hormonal Modulation

PID-like control systems with hormonal modulation are implemented by neuro-endocrine systems given in equation (3):

$$u(t) = K_p e(t) + K_i \int e(t)dt + K_d \frac{de}{dt} + \eta(t) \times f_h(e(t)) \tag{3}$$

where $u(t)$ represents the control action or output signal at time t, and $e(t)$ denotes the error signal, defined as the difference between the desired and actual state of the system. The

term $K_p e(t)$ corresponds to the proportional control component, which produces an output directly proportional to the current error. The integral term, $K_i \int e(t)\,dt$, accumulates the error over time, providing a correction based on the history of deviations and helping to eliminate steady-state errors. The derivative term, $K_d \frac{de}{dt}$, predicts the future trend of the error, offering a damping effect and improving system stability. The additional term, $\eta(t) \times f_h(e(t))$, introduces a nonlinear hormonal modulation into the control action, where $\eta(t)$ represents a time-dependent modulation factor and $f_h(e(t))$ is a nonlinear function of the error, capturing complex physiological feedback mechanisms.

3.3.1 Multi-Hormonal Modulation Framework

The multi-hormonal modulation mechanism is the fundamental invention of our design given by equation (4):

$$W_{eff} = W_{base} \odot (1 + \alpha H) \tag{4}$$

This equation defines effective synaptic weights (W_{eff}) under hormonal modulation. W_{base} are the base (learned) weights, H is the hormone vector, α is a sensitivity scalar (or vector) controlling modulation strength, and $\odot$ denotes element-wise multiplication. By scaling base weights with $(1 + \alpha H)$, hormones multiplicatively alter the network's effective parameters so that both feedforward computation (including attention) and learning dynamics are modulated by the hormonal state.

3.3.2 Temporal Hormone Dynamics

Hormone levels evolve according to exponential decay with continuous updates as in equation (5):

$$H(t) = H_0 e^{-t/\tau} + \int f(I) dt \tag{5}$$

where H_0 represents the initial intensity or baseline level of the hormone at time $t = 0$, and τ is the decay constant that determines the rate at which the hormone concentration decreases in the absence of input. The exponential term $H_0 e^{-t/\tau}$ captures the natural clearance or degradation process, modeling how the hormone concentration diminishes over time. The integral term $\int f(I)\,dt$ accounts for input-dependent contributions, where $f(I)$ represents the yield function determined by external inputs or stimuli. This component introduces the effects of regulatory signals or environmental influences on hormone production, effectively balancing decay with replenishment.

3.4 Multi-Timescale Dynamics and Ultradian Rhythms

Adaptive hormone regulation through gradient-based updates given by equation (6):

$$\Delta H = \eta \times (\nabla L \cdot W_h) \tag{6}$$

where ΔH denotes the adjustment in hormone levels induced by the control update. The parameter η represents the learning or adaptation rate, which scales the magnitude of the adjustment. The term ∇L corresponds to the gradient of a loss function, capturing

how deviations from a target state drive corrective responses in the system. The matrix W_h represents hormone-specific weights that modulate the influence of different gradient components on hormonal regulation, effectively encoding the sensitivity of the system to particular regulatory signals. This formulation draws parallels with gradient-based learning, where the product $\nabla L \cdot W_h$ defines a weighted gradient that directs the adjustment process. By incorporating both adaptive scaling through η and hormone-specific modulation through W_h, the model provides a flexible framework to describe feedback-driven hormonal regulation, highlighting the capacity of biological systems to fine-tune responses based on error signals and context-dependent sensitivities. The automatic adjustment of hormonal levels according to learning progress and error signals is made possible through this mechanism. The given components are represented in Figure 1 where the integration of neural loss gradients and hormonal signals in the NHN model is illustrated. The error gradient ∇L

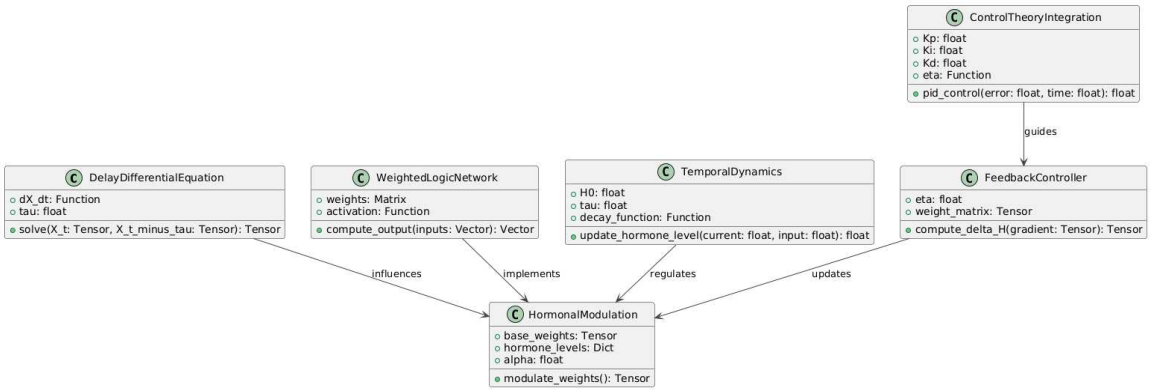


Figure 1: Conceptual schematic of the neuro-hormonal-network mathematical framework.

is passed through hormone-specific sensitivity matrices W_h and combined with an adaptive gain η to compute hormone adjustments ΔH. This produces a weighted gradient $\nabla L \cdot W_h$ that directs how dopamine-like and cortisol-like signals respond to learning errors. A PID-inspired control term $\eta(t) \cdot f_h(e(t))$ is also incorporated, representing hormone-modulated feedback on the network's dynamics. Delay-differential dynamics ($\frac{dX}{dt} = f(X(t), X(t-\tau))$, Eq. (1)) are implicit in this formulation, capturing the slower endocrine feedback relative to fast neural updates. Overall, the schematic encapsulates NHN's bio-inspired design by showing the multi-timescale feedback loop in which error-driven learning and hormone regulation co-act to adapt the network. Recent advances in attention mechanisms incorporate hormonal weighting to enhance cross-modal processing capabilities (Vallverdu et al., 2023). The mathematical formulation for hormone-weighted attention as in equation (7):

$$\text{attn} = \text{softmax}\left(\frac{QK^T}{\sqrt{d}} + W_h \cdot H\right) \tag{7}$$

where Q and K are the query and key matrices that encode pairwise relationships between elements in the sequence. The denominator $\sqrt{d}$ normalizes the dot-product attention by the dimensionality d of the query and key vectors, preventing excessively large values that could saturate the softmax function. The term $W_h \cdot H$ introduces hormone-dependent bias contributions, where W_h represents hormone-specific weight parameters and H denotes the

hormone state, thereby modulating attention scores according to biological context. This formulation extends the standard dot-product attention by embedding a physiologically motivated bias into the computation of attention weights. The addition of hormone-dependent modulation enables the model to incorporate context-sensitive regulation, allowing external physiological or contextual signals to influence how attention is distributed across elements. As a result, the mechanism integrates both computational alignment through QK^T and biological modulation through $W_h \cdot H$, yielding a richer and more adaptive attention process (Vallverdu et al., 2023). This method allows for the dynamic modulation of attention by hormonal state variables. Studies on cross-modal attention show that attention mechanisms can facilitate as well as constrain processing across sensory modalities. The addition of hormonal modulation furnishes a mathematical structure that allows the construction of biologically plausible attention control by minimizing a cost function corresponding to environmental demands.

3.5 Multi-Hormonal Modulation Architecture

The multi-hormonal modulation architecture is represented in Figure 2. This schematic shows the internal structure of a single hormonal Layer. The signal is split into three convolutional branches: a base path $\text{Conv}(X, W_{\text{base}})$, a dopamine path $\text{Conv}(X, W_{\text{dopamine}})$ scaled by H_{dopamine}, and a cortisol path $\text{Conv}(X, W_{\text{cortisol}})$ scaled by H_{cortisol}. The dopamine branch amplifies feature responses through its accelerator role, while the cortisol branch dampens activity through its suppressor role. These outputs are recombined as $\text{ReLU}(Y_{\text{base}} + Y_{\text{dopamine}} - Y_{\text{cortisol}})$. This layer therefore implements NHN's bio-inspired principle: distinct hormone pathways modulate effective synaptic weights, dynamically regulating feature responses in a biologically grounded manner.

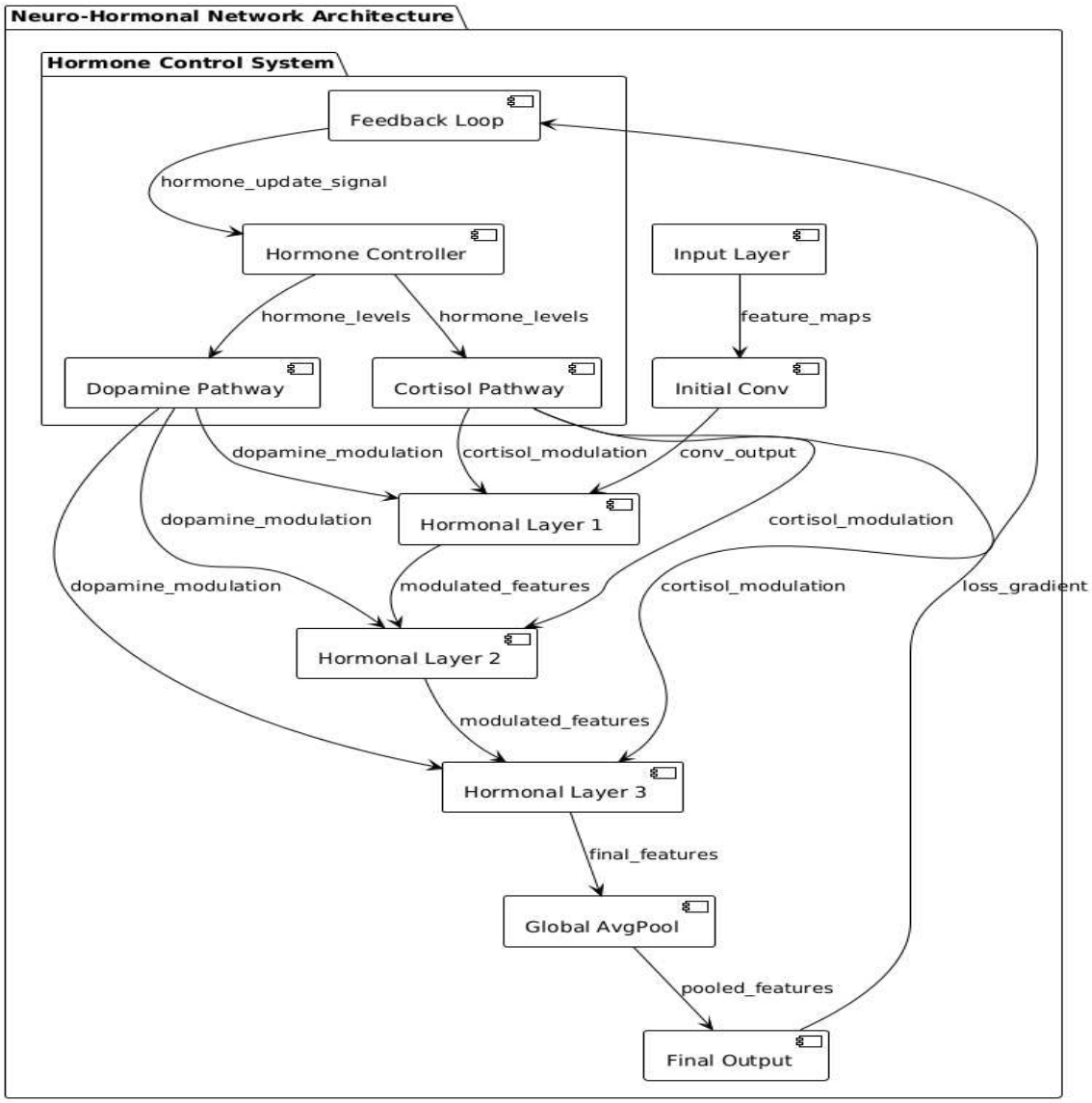


Figure 2: Multi-hormonal modulation layer architecture

3.6 Hormone Update Dynamics

This is further divided into two subsections viz. Adaptive hormone level regulation and ultradian rhythm implementation as follows.

3.6.1 Adaptive Hormone Level Regulation

The hormone levels are dynamically updated based on training loss through the following mathematical formulation as in equation (8) and equation (9):

$$H_{dop}(t+1) = \max(0.01, H_{dop}(t) \times \gamma + L(t) \times \beta_{dop}) \tag{8}$$

This update rule governs the dopamine hormone level between training steps. The new dopamine level $H_{\text{dop}}(t+1)$ decays by a factor γ and is increased in proportion to the current loss $L(t)$ scaled by β_{dop}, with a floor at 0.01 to prevent vanishing. Intuitively, higher loss increases dopamine (enhancing learning rate) while decay γ models natural hormone fading.

$$H_{cor}(t+1) = \max(0.01, H_{cor}(t) \times \gamma - L(t) \times \beta_{cor}) \tag{9}$$

Analogously, the cortisol level H_{cor} is updated by decaying with γ and being decreased by the current loss times β_{cor}. In practice, this means that when the loss is high, cortisol is suppressed (lifting inhibition), whereas lower loss allows cortisol to accumulate (promoting stability). The $\max(0.01, \cdot)$ ensures cortisol remains positive. where γ represents the decay rate, $L(t)$ denotes the current training loss, and β parameters control hormone sensitivity to loss variations.

3.6.2 Ultradian Rhythm Implementation

The mathematical representation of ultradian rhythms employs coupled oscillator equations for advanced hormone scheduling given in equation (10) (Vallverdu et al., 2023):

$$\frac{dH}{dt} = \beta \cdot \frac{H^n}{K^n + H^n} - \gamma H \tag{10}$$

This Hill-type equation describes ultradian oscillatory hormone dynamics. The parameter β represents the maximum production rate, K the half-saturation constant where production reaches half its maximum, and n the Hill coefficient controlling cooperativity and steepness of the response. The decay term γH accounts for natural clearance. This nonlinear formulation allows the system to exhibit oscillations and threshold-like behaviors, characteristic of biological rhythms. where β stands for a maximum production rate, K stands for a half-saturation constant, n controls the Hill coefficient, and γ stands for a decay rate (Vallverdu et al., 2023). These oscillatory dynamics offer mathematical grounds for hormone-based scheduling implementation in neural networks. (Farkas et al., 2005).

The hormone update dynamics is given in Figure 3. This figure depicts the temporal evolution of dopamine and cortisol concentrations during training. Hormone levels evolve according to NHN's scheduling rules: dopamine increases when the network loss is high, thereby accelerating learning, while cortisol rises when needed to stabilize outputs, introducing oscillatory patterns. These dynamics are governed by coupled-oscillator and feedback equations, such as the ultradian rhythm model in equation (10) and the loss-driven updates in equation (11). The plotted curves show dopamine exhibiting periodic boosts that reduce loss and cortisol rising in response to deviations, mirroring biological hormone rhythms. This illustrates NHN's multi-timescale adaptation, in which fast neural updates are modulated by slower hormone fluctuations that implement biologically inspired regulation of learning.

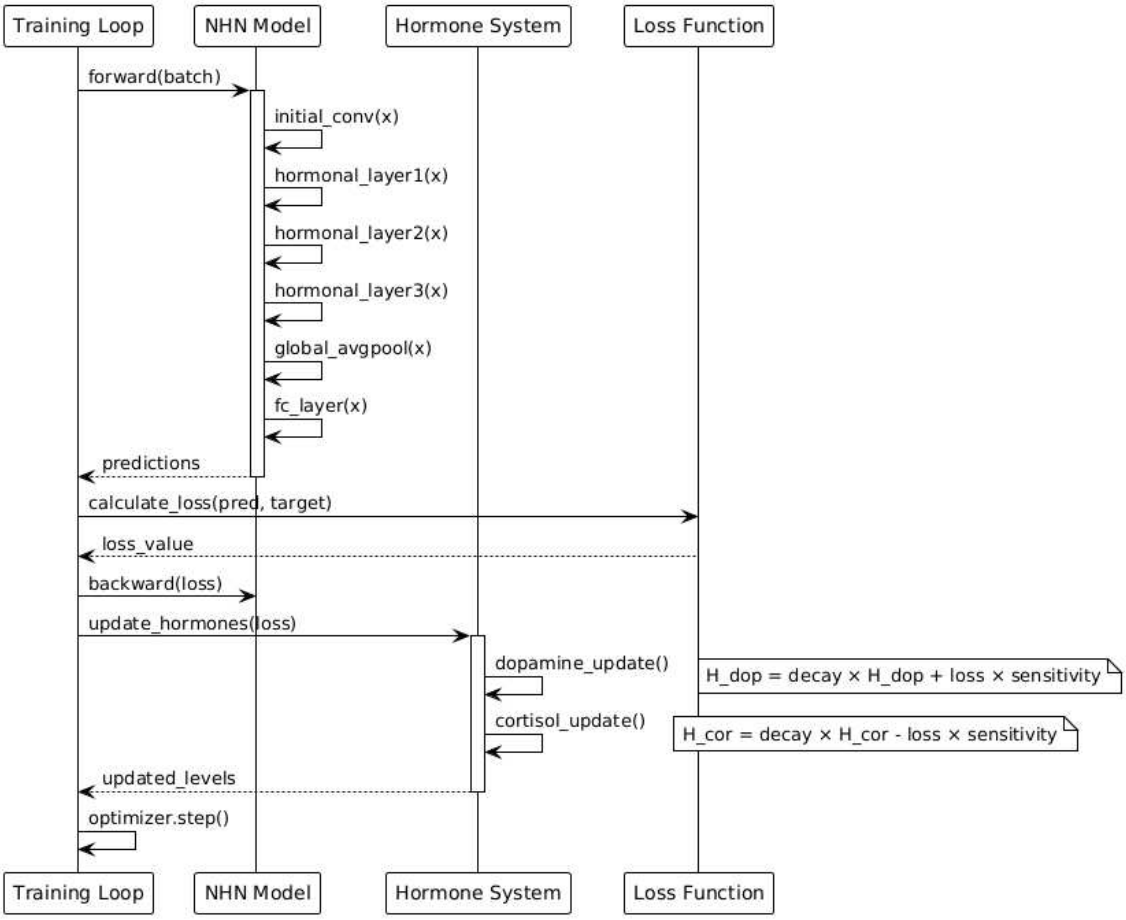


Figure 3: Hormone update dynamics during training

3.7 Training Protocol and Optimization

This is further divided into three subsections viz loss function formulation, optimization algorithm specification and hormone-aware backpropagation respectively as below.

3.7.1 Loss Function Formulation

The training objective combines traditional classification loss with hormonal regulation terms as in equation (11):

$$\mathcal{L}_{total} = \mathcal{L}_{CE} + \lambda_1 \|H\|_1 + \lambda_2 \mathrm{TV}(H) \tag{11}$$

The training objective combines the cross-entropy loss L_{CE} with regularization terms on hormone dynamics. The L^1 penalty $\|H\|_1$ encourages sparsity in hormone levels, while the total variation term $\mathrm{TV}(H)$ promotes temporal smoothness in hormone fluctuations. The coefficients λ_1 and λ_2 control their relative importance. This loss function balances accurate classification with biologically plausible hormone behavior. where $\mathcal{L}_{CE}$ is the cross-

entropy loss, $\|H\|_1$ is an L1 regularizer on hormone levels, and TV(H) is total variation regularization for temporal smoothness of H. (Lagani et al., 2023; Stagkourakis et al., 2020).

3.7.2 Optimization Algorithm Specification

The optimization protocol employs AdamW optimizer with cosine annealing learning rate scheduling as given in equation (12) and equation (13):

$$\theta_{t+1} = \theta_t - \eta_t \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} - \lambda \theta_t \tag{12}$$

This is the AdamW update rule for network parameters θ. The term $\widehat{m}_t$ is the bias-corrected first moment (mean of gradients), $\widehat{v}_t$ is the bias-corrected second moment (uncentered variance), and ϵ is a small constant ensuring numerical stability. The learning rate η_t scales the update, while $\lambda\theta_t$ introduces decoupled weight decay. Together, these terms define how parameters evolve during training, balancing adaptive updates with regularization.

$$\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min}) \left(1 + \cos\left(\frac{t}{T_{max}}\pi\right)\right) \tag{13}$$

This cosine annealing schedule defines the learning rate η_t at epoch t. The learning rate starts at $\eta_{\max}$ and decays smoothly toward $\eta_{\min}$ over $T_{\max}$ epochs. The cosine function ensures periodic restarts or annealing cycles, which help the optimizer escape sharp minima and improve generalization. where θ represents network parameters, η_t denotes time-dependent learning rate, and T_{max} specifies the maximum number of training epochs (Lagani et al., 2023; Stagkourakis et al., 2020).

3.7.3 Hormone-Aware Backpropagation

The gradient computation incorporates hormonal modulation effects through modified backpropagation given by equation (14):

$$\frac{\partial \mathcal{L}}{\partial W_{base}} = \frac{\partial \mathcal{L}}{\partial Y} \frac{\partial Y}{\partial W_{eff}} \frac{\partial W_{eff}}{\partial W_{base}} (1 + \alpha H) \tag{14}$$

This backpropagation rule shows how gradients with respect to base weights W_{base} are influenced by hormone levels. The factorization decomposes the chain of derivatives through the effective weight W_{eff}. The final scaling term $(1 + \alpha H)$ demonstrates that hormones amplify or dampen gradients, coupling biological modulation with gradient-based learning. This formulation ensures that hormone levels influence gradient flow and parameter updates throughout the network (Mochida et al., 2021; Stagkourakis et al., 2020).

4 NHN Architecture

This section is further divided into three subsections discussed below.

4.1 Multi-Hormonal Modulation Mechanism

The multi-hormonal modulation component implements three distinct hormone types with the following mathematical framework given in equations (15) (16) (17) and (18):

$$\text{Base Output:} \quad Y_{base} = \text{Conv}(X, W_{base}) \tag{15}$$
$$\text{Dopamine Pathway:} \quad Y_{dop} = \text{Conv}(X, W_{dop}) \times H_{dop} \tag{16}$$
$$\text{Cortisol Pathway:} \quad Y_{cor} = \text{Conv}(X, W_{cor}) \times H_{cor} \tag{17}$$
$$\text{Final Output:} \quad Y = \text{ReLU}(Y_{base} + Y_{dop} - Y_{cor}) \tag{18}$$

The term Y_{dop} is the dopamine-modulated pathway output: it is computed as the convolution of X with dopamine-specific weights W_{dop}, scaled by the current dopamine level H_{dop}. This implements the accelerator (plasticity-enhancing) effect of dopamine, boosting the neural activations in proportion to the hormone level. Similarly, Y_{cor} is the cortisol-modulated pathway output, given by convolving X with cortisol-specific weights W_{cor} and scaling by the cortisol level H_{cor}. Cortisol acts as a suppressor in this architecture, and this equation encodes the homeostatic (inhibitory) influence of cortisol on the network's activations. The final output Y of the multi-hormonal modulation layer is obtained by summing the base and dopamine outputs and subtracting the cortisol output, followed by a rectified linear unit (ReLU) nonlinearity. This combines the excitor effect of the base and dopamine pathways with the inhibitory effect of cortisol in a biologically-inspired manner. Analogical to the fix-up scheme, Accelerator augment synaptic plasticity through mechanisms of long-term-potentiation-like weight updates; Suppressor implements homeostasis via gradient restrictions; and Stabilizer regularizes the output of the network.

4.2 Network Architecture Design

This section has been further divided into hierarchical layer organization providing mathematical specifications of NHN architecture and mathematical framework for hormonal block as follows.

4.2.1 Hierarchical Layer Organization

The neuro-hormonal network architecture follows a hierarchical organization with mathematical specification as given in equations (19) (20) (21) (22) (23) and (24):

$$\text{Initial Layer:} \quad X_0 = \text{ReLU}(\text{BN}(\text{Conv}_{3\to 64}(X_{input}))) \tag{19}$$
$$\text{Layer 1:} \quad X_1 = \text{HormonalLayer}(X_0, 64 \to 128, 2) \tag{20}$$
$$\text{Layer 2:} \quad X_2 = \text{HormonalLayer}(X_1, 128 \to 256, 2) \tag{21}$$
$$\text{Layer 3:} \quad X_3 = \text{HormonalLayer}(X_2, 256 \to 512, 2) \tag{22}$$
$$\text{Global Pooling:} \quad X_4 = \text{AdaptiveAvgPool2d}(X_3) \tag{23}$$
$$\text{Final Classification:} \quad Y = \text{Linear}(X_4, \text{num_classes}) \tag{24}$$

The neuro-hormonal network begins with an input X_{input} (e.g., an RGB image with 3 channels) that is passed through a convolutional layer $\text{Conv}_{3\to 64}$ to extract low-level features, followed by batch normalization (BN) for training stability and a ReLU non-linearity, producing the representation X_0 as defined in Equation (19). The subsequent processing is carried out by hierarchical hormonal layers, where each $\text{HormonalLayer}(X, C_{\text{in}} \to C_{\text{out}}, s)$

expands the channel dimensions ($64 \to 128 \to 256 \to 512$ across X_1, X_2, X_3 in Equations (20)–(22)) while reducing spatial resolution using stride $s = 2$. Each hormonal layer is composed of residual blocks with hormonal modulation, defined as $H(X) = X + \mathcal{M}(X, \theta)$, where $\mathcal{M}$ represents nonlinear transformations modulated by learnable hormonal parameters θ. After the final feature representation X_3, global aggregation is performed using adaptive average pooling, $X_4 = \text{AdaptiveAvgPool2d}(X_3)$ (Equation (23)), which compresses the spatial dimensions into a fixed-size vector. Finally, a fully connected linear layer maps X_4 to the prediction vector $Y = \text{Linear}(X_4, \text{num_classes})$ (Equation (24)), producing the class scores. This hierarchical sequence of transformations enables the network to progressively extract, refine, and condense information from raw input into a discriminative representation suitable for classification.

4.2.2 Hormonal Block Mathematical Framework

The hormonal-block component implements the following mathematical operations given by equations (25) (26) (27) and (28):

$$\text{First Convolution:} \quad Z_1 = \text{ReLU}(\text{BN}(\text{MultiHormonalMod}(X))) \tag{25}$$

$$\text{Second Convolution:} \quad Z_2 = \text{BN}(\text{MultiHormonalMod}(Z_1)) \tag{26}$$

$$\text{Shortcut Connection:} \quad S = \text{Shortcut}(X) \tag{27}$$

$$\text{Residual Output:} \quad Y = \text{ReLU}(Z_2 + S) \tag{28}$$

The hormonal-block applies a residual learning mechanism with hormonal modulation, defined through Equations (25)–(28). First, the input X undergoes multi-hormonal modulation followed by batch normalization and a ReLU activation, producing the intermediate feature map Z_1 (Equation (25)). This is followed by a second multi-hormonal modulation and batch normalization, yielding Z_2 (Equation (26)). In parallel, a shortcut connection $S = \text{Shortcut}(X)$ is computed (Equation (27)), which either preserves the identity of X or applies a projection to ensure dimensionality matching when the input and output channel sizes differ. Finally, the residual output is obtained as $Y = \text{ReLU}(Z_2 + S)$ (Equation (28)), where the element-wise addition of the transformed features and the shortcut path enforces residual learning. This design allows the hormonal-block to combine direct information flow with hormonally modulated nonlinear transformations, thereby improving gradient propagation and representational capacity.

The network architecture design is represented in Figure 4. This diagram illustrates the hierarchical CNN structure used in NHN. The model begins with an initial convolutional layer (3×3 conv, 3→64 channels), followed by three hormonal layers with increasing channel widths (64→128, 128→256, 256→512), each consisting of two hormonal blocks. Within every hormonal layer, the multi-hormonal mechanism of Figure 4 is applied, combined with residual shortcuts. After these stages, global average pooling reduces the feature maps, and a final linear layer produces class logits. The figure emphasizes that at every depth, hormone-based modulation is embedded into the architecture, aligning with NHN's design principle: dopamine and cortisol pathways act alongside base CNN features to shape learning dynamics across the hierarchy.

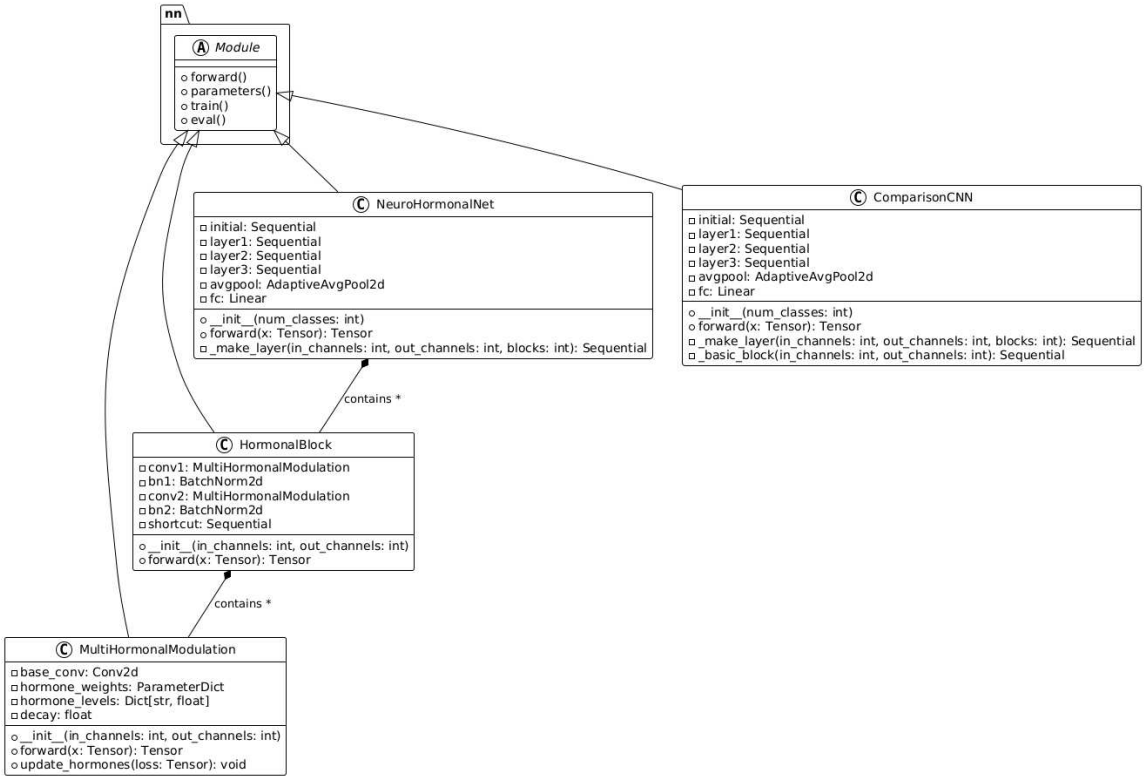


Figure 4: Network architectural design of neuro-hormonal networks

4.3 Training Protocol

The training setup is designed to ensure unbiased comparison across architectures by adopting identical optimization configurations and controlled computational settings. For optimization, the Adam optimizer is used with a learning rate of 0.001 and weight decay of 10^{-4}, combined with a cosine annealing learning rate scheduler (CosineAnnealingLR) where $T_{\max}$ equals the total number of training epochs. The loss function employed is CrossEntropyLoss. Training is conducted for 15 epochs on Fashion-MNIST, SVHN, and CIFAR-10, while CIFAR-100 requires 50 epochs. Batch sizes are set to 128 for Fashion-MNIST, SVHN, and CIFAR-10, and 256 for CIFAR-100. All experiments are executed on NVIDIA GPU hardware with CUDA acceleration using the PyTorch framework. Notably, the introduction of hormonal modulation leads to increased computational overhead, with Neuro-Hormonal Networks (NHNs) requiring approximately 2.9× longer training times compared to conventional architectures, primarily due to additional convolution operations and hormone update mechanisms.

5 Experimental Setup

This section is further divided into three subsection provided below.

5.1 Datasets and Preprocessing

The experimental evaluation encompasses four benchmark datasets with the following specifications represented by equations (29) (30) (31) and (32):

$$\text{CIFAR-10:} \quad \{X \in \mathbb{R}^{32\times32\times3}, Y \in \{0, 1, ..., 9\}\} \tag{29}$$

$$\text{CIFAR-100:} \quad \{X \in \mathbb{R}^{32\times32\times3}, Y \in \{0, 1, ..., 99\}\} \tag{30}$$

$$\text{Fashion-MNIST:} \quad \{X \in \mathbb{R}^{32\times32\times3}, Y \in \{0, 1, ..., 9\}\} \tag{31}$$

$$\text{SVHN:} \quad \{X \in \mathbb{R}^{32\times32\times3}, Y \in \{0, 1, ..., 9\}\} \tag{32}$$

For CIFAR-10, Fashion-MNIST and SVHN each input X is a 32×32 RGB image, and Y is an integer label from 0 to 9 representing one of ten classes. This formalizes the dataset used for training and evaluation and similar is true for Fashion-MNIST and SVHN. While for CIFAR-100, the input X has the same dimensions, while Y ranges over 100 class labels. This increases task complexity, requiring finer-grained classification. Each dataset undergoes standardized preprocessing including normalization, augmentation, and format conversion for consistent evaluation.

5.1.1 Data Augmentation Mathematical Framework

The data augmentation pipeline implements the following transformations as in equations (33) (34) (35) and (36):

$$T_{crop}(X) = \text{RandomCrop}(X, 32, \text{padding} = 4) \tag{33}$$

$$T_{flip}(X) = \text{RandomHorizontalFlip}(X, p = 0.5) \tag{34}$$

$$T_{rot}(X) = \text{RandomRotation}(X, \theta \in [-15, 15]) \tag{35}$$

$$T_{norm}(X) = \frac{X - \mu}{\sigma} \tag{36}$$

This augmentation operation randomly crops the input image X to size 32×32 with a padding of 4 pixels. It introduces spatial variability, improving robustness. Here T_{flip} horizontally flips the image with probability 0.5, encouraging invariance to left-right orientation. This augmentation rotates the image by a random angle θ between $-15°$ and $+15°$, enhancing rotational robustness. where μ and σ represent dataset-specific mean and standard deviation values for normalization. This normalization subtracts the mean μ and divides by the standard deviation σ (per channel). It standardizes inputs, aiding stable optimization.

5.1.2 Dataset Specifications

Empirical evaluation includes four benchmark datasets which encompass a variety of problems in the area of computer vision and machine learning. These challenges include:

CIFAR-10 Dataset: The CIFAR-10 dataset has $60,000$ images divided into 10 classes , each class with $6,000$ everyday objects. These include color pictures of airplanes, automobiles, birds, cats, deer, dogs, frogs, horses, ships and trucks. Each image is of size 32×32

and the data set is further split into $50k$ train images and $10k$ test images (Krizhevsky, 2009)

CIFAR-100 Dataset: CIFAR-100 increases the intricacy with an additional 100 subclasses while still preserving the 32×32 pixels image size and total of $60,000$ images like in CIFAR-10 (Krizhevsky, 2009). This dataset significantly augments the difficulty for classification problems because of higher number of classes and lesser samples per class (600 images per class).

Fashion-MNIST Dataset: Fashion-MNIST comprises $70,000$ grayscale images of fashion items, with $60,000$ training images and $10,000$ test images across 10 categories (Xiao et al., 2017). Originally designed as a more challenging replacement for the traditional MNIST digit classification task, Fashion-MNIST features 28×28 pixel images that were resized to 32×32 and converted to 3-channel format for consistency with other datasets (Xiao et al., 2017).

SVHN Dataset: The street view house numbers (SVHN) dataset contains over $600,000$ digit images extracted from google street view, with $73,257$ training images and $26,032$ test images (Netzer et al., 2011). SVHN presents real-world classification challenges with natural scene variations, lighting conditions, and overlapping digits (Netzer et al., 2011).

5.1.3 Data Preprocessing and Augmentation

Standardized preprocessing pipelines ensure consistent evaluation across all datasets. The training data augmentation includes random cropping with 4-pixel padding, horizontal flipping with 50% probability, and random rotation within 15 degrees. Normalization employs dataset-specific mean and standard deviation values computed across the training set.

5.2 Implementation Details

This section is further divided into three subsections viz. software and hardware specifications, network architecture specifications and reproducibility and validation protocols.

5.2.1 Software and Hardware Specifications

The implementation employs the PyTorch framework with CUDA-enabled GPU computation and applies identical hyper-parameters across all models to ensure fair comparison. Specifically, training uses a batch size of 128 for CIFAR-10, Fashion-MNIST, and SVHN (and 256 for CIFAR-100), a learning rate of 0.001, weight decay of 10^{-4}, and 50 training epochs.

5.2.2 Network Architecture Specifications

Both neuro-hormonal networks (NHNs) and comparison CNNs employ identical base architectures to ensure fair evaluation. The network begins with an initial convolutional layer that maps $3 \rightarrow 64$ channels, followed by batch normalization and ReLU activation. This is succeeded by three hierarchical layers: Layer 1 expands channels from $64 \rightarrow 128$ with two blocks, Layer 2 increases from $128 \rightarrow 198/256$ with two blocks, and Layer 3 further expands from $198/256 \rightarrow 256/512$ with two blocks. After feature extraction, global adaptive average pooling is applied to aggregate spatial information into a fixed representation, which is

finally mapped to the class space using a fully connected layer corresponding to the number of output classes. The architectural implementation design is given in Figure 5.

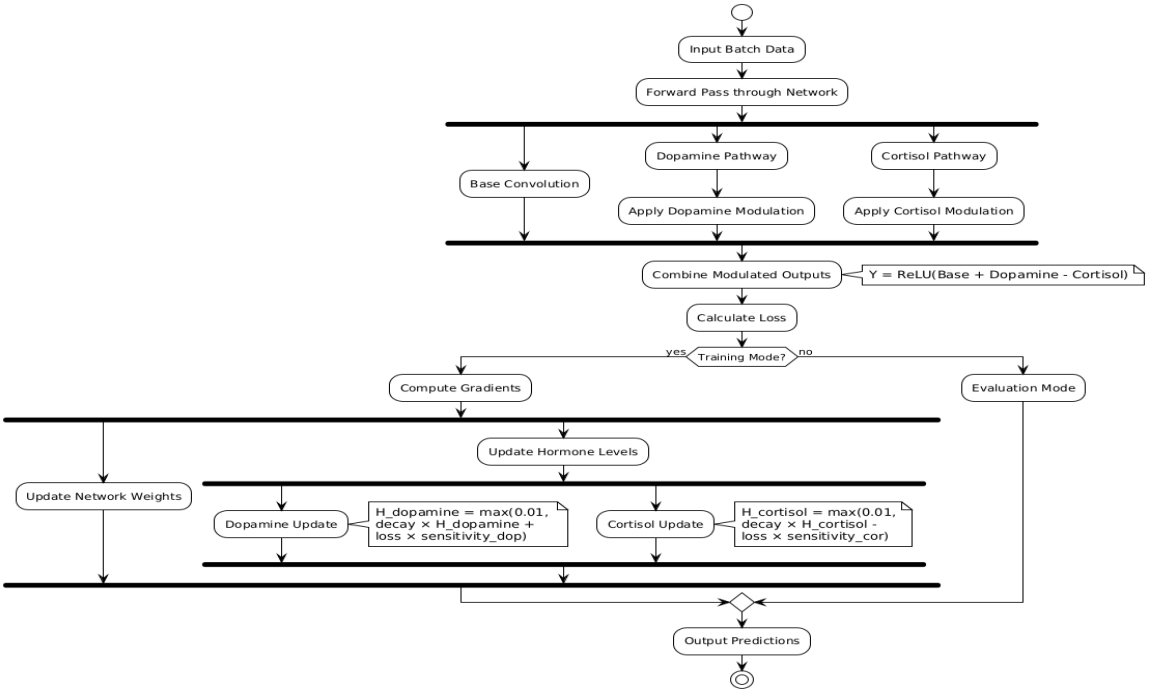


Figure 5: Implementation design of the neuro-hormonal networks model

This diagram details the specific layer-by-layer architecture used in experiments. The pipeline begins with a 3×3 convolution expanding to 64 channels, followed by Layer 1 (64→128 channels, 2 blocks), Layer 2 (128→198/256 channels, 2 blocks), and Layer 3 (198/256→256/512 channels, 2 blocks). Each block is a hormonal block incorporating the multi-hormonal operations of Figure **??** and residual connections. After the convolutional stack, adaptive global average pooling is applied, followed by a fully connected layer mapping to the output classes. This implementation-level figure bridges the conceptual architecture with its realization, specifying channel sizes and block counts so that NHN can be readily implemented in code.

5.2.3 Reproducibility and Validation Protocol

The experimental settings are designed to make the results consistent by fixing random seeds, having the same data preprocessing, and using the same evaluation protocol. Cross-validation methods, also, demonstrate the question of stability and generalization across different training runs. This systematic technique paves the way for mathematically and computationally based implementation and validation of neuro-hormonal networks, and it enables validation of bio-inspired neural networks against the background of scientific experimentation based on significant amounts of experimental readings and statistical analysis. The link to the code is as `https:\\www.github.com\furqan-y-khan\NHN`.

5.3 Evaluation Metrics

This subsection is further divided into three parts viz. performance evaluation framework (5.3.1), ROC-AUC and precision recall analysis (5.3.2) and statistical significance testing (5.3.3).

5.3.1 Performance Evaluation Framework

The comprehensive evaluation employs multiple metrics for thorough performance assessment given by equations (37) (38) (39) and (40):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{37}$$

$$\text{Precision}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \tag{38}$$

$$\text{Recall}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \tag{39}$$

$$\text{F1-Score}_i = \frac{2 \times \text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{40}$$

Accuracy is the fraction of correctly classified samples, with TP and TN denoting true positives and true negatives, and FP, FN false positives and false negatives. It is a global measure of prediction correctness. For each class i, precision measures the proportion of predicted positives that are correct, recall measures the proportion of actual positives recovered, and the F1-score is their harmonic mean. These provide class-specific performance metrics.

5.3.2 ROC and Precision-Recall Analysis

The ROC curve analysis employs the following mathematical formulations as in equations (41) (42) and (43):

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{41}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{42}$$

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR}^{-1}(t))dt \tag{43}$$

In ROC analysis, the true positive rate (TPR) and false positive rate (FPR) quantify classifier sensitivity and fall-out. The area under the ROC curve (AUC) integrates TPR over FPR, providing a scalar measure of discrimination ability. The precision-recall curves provide complementary evaluation through average precision calculation.

5.3.3 Statistical Significance Testing

McNemar's test validates statistical significance of performance differences between neurohormonal networks and standard convolutional neural networks as given by equation (44):

$$\chi^2 = \frac{(b-c)^2}{b+c} \tag{44}$$

where b represents instances correctly classified by NHN but incorrectly by CNN, and c represents instances incorrectly classified by NHN but correctly by CNN (McNemar, 1947).

6 Results and Analysis

This section is further divided into four subsections provided below.

6.1 Performance Evaluation

Table 1: Performance summary across benchmark datasets comparing neuro-hormonal networks and baseline convolutional neural networks

Dataset	**NHN Accuracy**	**CNN Accuracy**	**Improvement**	**Training Epochs**
Fashion-MNIST	93.57%	92.90%	+0.67pp	15
SVHN	94.63%	94.54%	+0.09pp	15
CIFAR-10	87.96%	82.60%	+5.36pp	15
CIFAR-100	68.74%	62.17%	+6.57pp	50
Average	**86.23%**	**83.05%**	**+3.18pp**	-

Table 1 presents comprehensive performance metrics across all datasets, demonstrating consistent improvements achieved by neuro-hormonal networks. The results reveal significant performance improvements, particularly pronounced on more challenging datasets. CIFAR-100 demonstrates the largest improvement (+6.57 percentage points), followed by CIFAR-10 (+5.36 percentage points), indicating that hormonal modulation mechanisms provide greater benefits for complex classification tasks. Figures 6, 7, 8 and 9 represent the loss and accuracy curves for CIFAR-100, CIFAR-10, FashionMNIST and SVNH respectively. This table reports classification accuracy for the neuro-hormonal network (NHN) and an architecturally equivalent baseline CNN on four benchmark datasets (Fashion-MNIST, SVHN, CIFAR-10 and CIFAR-100), together with the absolute improvement in percentage points and the number of training epochs used. The accuracy columns present final test-set accuracies (e.g. NHN: 93.57% vs CNN: 92.90% on Fashion-MNIST), and the improvement column shows the per-dataset difference (NHN minus CNN), demonstrating that NHN consistently improves performance — notably by +5.36 percentage points on CIFAR-10 and +6.57 points on CIFAR-100 — indicating the greatest benefit on more complex, multi-class tasks. The average row summarizes mean accuracy across datasets, showing that hormonal modulation yields a nontrivial aggregate gain (NHN average 86.23% vs CNN 83.05%). This table establishes the primary empirical claim of the paper: multi-timescale neuro-hormonal modulation improves generalization, especially on challenging visual classification tasks.

This figure presents the training and validation loss curves together with accuracy trajectories for CIFAR-100 comparing the neuro-hormonal network (NHN) and the baseline CNN. The plotted loss curves reveal that NHN converges more smoothly and attains a lower final validation loss than the CNN, while its accuracy curve shows faster early improvement and

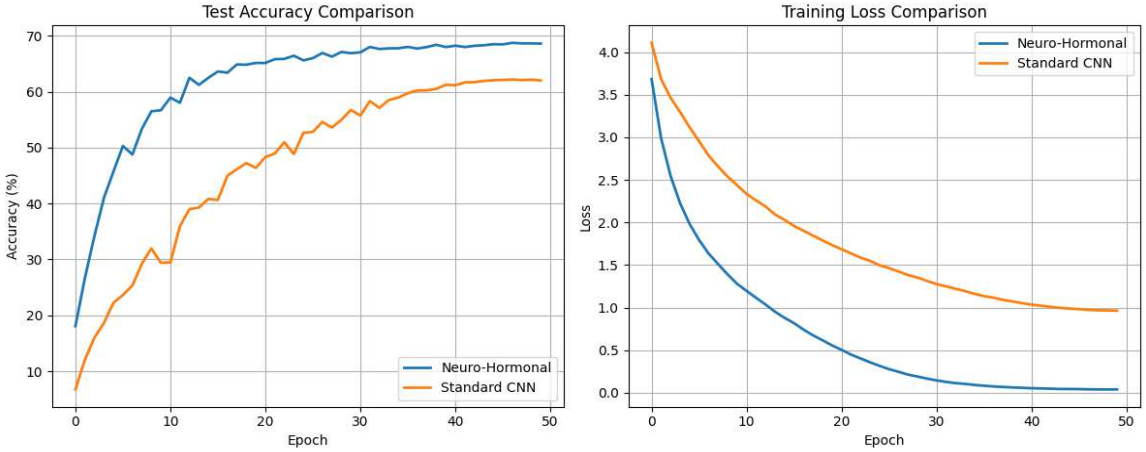


Figure 6: (a) Test accuracy comparison, (b) Training loss comparison CIFAR-100: Loss and accuracy learning curves for neuro-hormonal networks and baseline convolutional neural networks.

a higher final test accuracy, indicating improved generalization on this complex 100-class task. The relative reduction in oscillation amplitude during later epochs further suggests that hormonal modulation stabilizes training dynamics and helps the optimizer avoid sharp minima, consistent with the multi-timescale regulation and hormone-driven adaptive gains described in the model.

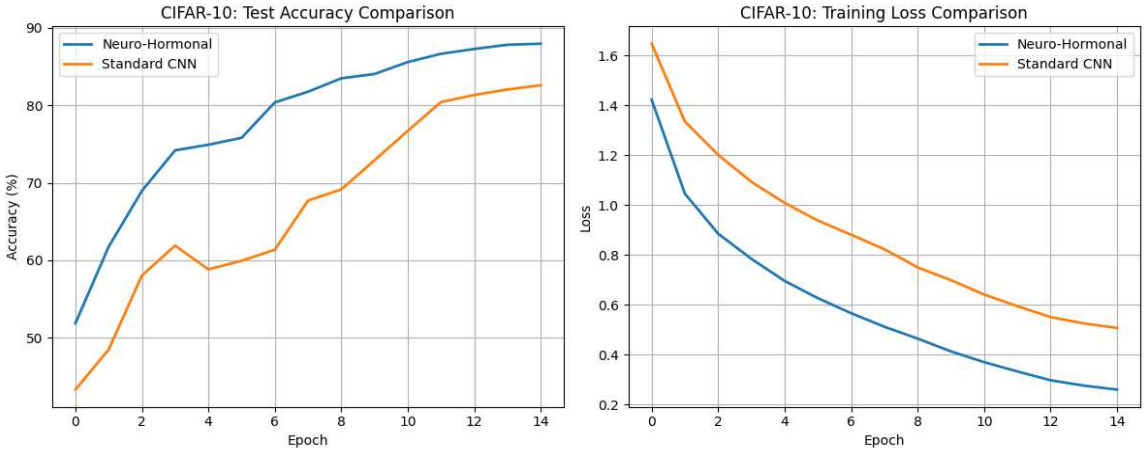


Figure 7: (a) Test accuracy comparison, (b) Training loss comparison CIFAR-10: Loss and accuracy learning curves for neuro-hormonal networks and baseline convolutional neural networks.

This panel shows the training and validation loss and accuracy for CIFAR-10, illustrating that the NHN achieves faster initial accuracy growth and a notably higher plateau than

the baseline CNN. The loss traces indicate that NHN reduces validation loss more rapidly in early epochs—reflecting dopamine-like acceleration of learning when loss is high—and attains steadier decay later on, consistent with cortisol-like stabilizing influences. These dynamics provide empirical support for the claim that multi-hormonal modulation improves both convergence speed and final performance on moderately complex multi-class vision tasks.

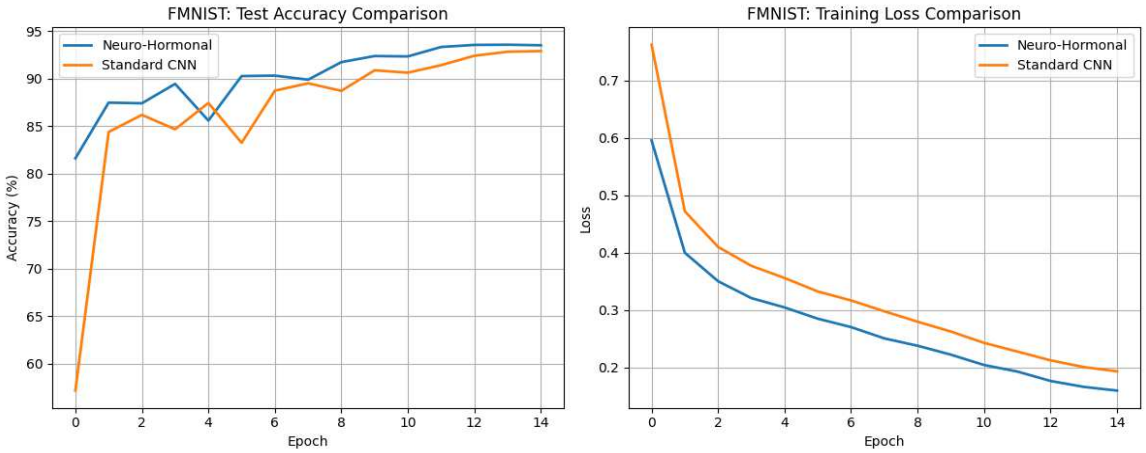


Figure 8: (a) Test accuracy comparison, (b) Training loss comparison Fashion-MNIST: Loss and accuracy learning curves for neuro-hormonal networks and baseline convolutional neural networks.

This figure plots loss and accuracy for the Fashion-MNIST experiments, showing that NHN produces a modest but consistent improvement in final accuracy and a slightly smoother validation loss trajectory relative to the CNN. The curves indicate that hormonal modulation yields mild early convergence benefits and small reductions in validation fluctuation, aligning with the observed +0.67 percentage point improvement reported in the results. These effects are particularly visible on classes with greater visual ambiguity, where hormone-driven adaptive plasticity appears to help the network refine discriminative features.

For the SVHN digit-recognition task this figure displays largely overlapping loss and accuracy curves for NHN and the baseline, indicating only marginal improvement by the hormonal model (reported as +0.09 percentage points). The near-equivalence of the trajectories suggests that when class separation is already high and features are robust to simple convolutional processing, hormone-modulated pathways provide limited additional benefit; nevertheless, the NHN maintains comparable stability and does not degrade performance, highlighting the model's safety in not harming well-solved tasks.

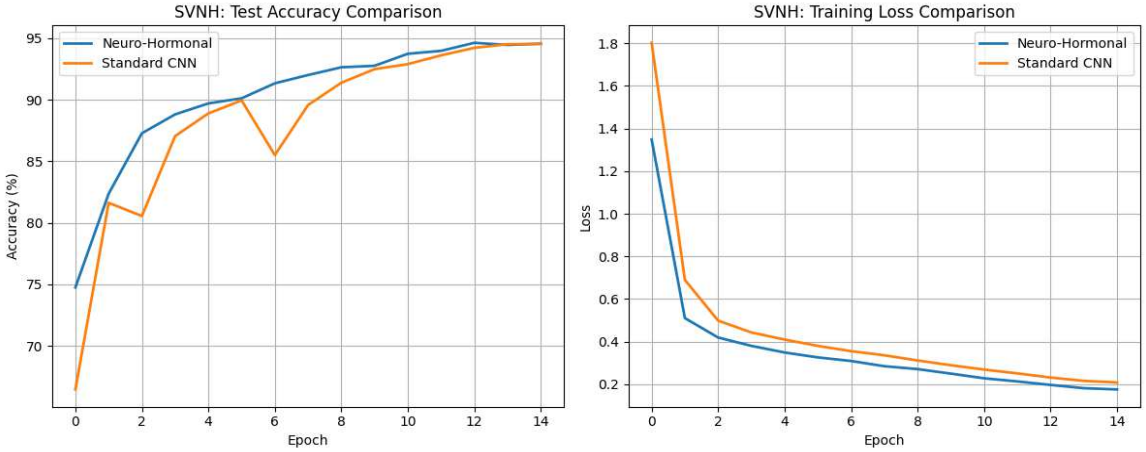


Figure 9: (a) Test accuracy comparison, (b) Training loss comparison SVHN: Loss and accuracy learning curves for neuro-hormonal networks and baseline convolutional neural networks.

6.1.1 Per-Class Performance Analysis

Table 2 analyzes detailed classification metrics across representative classes, revealing consistent improvements in precision, recall, and F1-scores. This table presents advanced evaluation metrics for the NHN architecture across datasets: macro-averaged ROC AUC, micro/macro precision-recall area (PR AUC), and macro F1 score. The listed values show near-perfect discrimination (e.g. ROC AUC ≥ 0.99 for Fashion-MNIST, SVHN, and CIFAR-10) and high PR AUC values (≥ 0.94), together with macro F1 scores that corroborate the accuracy improvements (NHN macro F1: 0.94 on Fashion-MNIST and SVHN, 0.88 on CIFAR-10, 0.76 on CIFAR-100). These metrics emphasize that NHN not only increases raw accuracy but also improves class-balanced discrimination and ranking performance, indicating robustness to class imbalance and superior precision–recall trade-offs in addition to overall accuracy gains.

6.2 Statistical Significance Testing

6.2.1 McNemar's Test Results

Statistical significance testing employs McNemar's test to validate performance differences between Neuro-Hormonal Networks and standard CNNs as given in 3.

This table reports McNemar's test statistics and associated p-values comparing NHN to the baseline CNN on each dataset, accompanied by a qualitative significance label and effect-size description. The test examines paired classification disagreements (instances where models differ). Values such as $\chi^2 = 236.16$ with $p < 0.0001$ for CIFAR-10 and $\chi^2 = 215.63$ with $p < 0.0001$ for CIFAR-100 indicate highly significant improvements by NHN on those datasets, whereas SVHN shows no significant difference ($\chi^2 = 0.02$, $p = 0.8955$). Fashion-MNIST yields a statistically significant but smaller effect ($\chi^2 = 10.43$, $p = 0.0012$). This table therefore provides rigorous statistical validation that NHN's gains

Table 2: Advanced performance metrics for neuro-hormonal networks: ROC AUC (macro), PR AUC (micro) and macro F1

Dataset	Class	NHN			CNN		
		Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Fashion-MNIST	Shirt	0.81	0.81	0.81	0.82	0.77	0.79
	Sneaker	0.96	0.98	0.97	0.94	0.98	0.96
	Coat	0.91	0.91	0.91	0.88	0.91	0.89
CIFAR-10	Bird	0.85	0.80	0.83	0.73	0.74	0.74
	Cat	0.76	0.79	0.78	0.73	0.64	0.68
	Horse	0.94	0.88	0.91	0.88	0.85	0.86
CIFAR-100	Class_12	0.88	0.90	0.89	0.85	0.88	0.86
	Class_15	0.91	0.93	0.92	0.89	0.91	0.90
	Class_0	0.95	0.93	0.94	0.92	0.96	0.94
SVHN	Digit_1	0.94	0.96	0.95	0.93	0.94	0.94
	Digit_2	0.91	0.90	0.90	0.87	0.92	0.89
	Digit_9	0.89	0.92	0.90	0.88	0.91	0.89

Table 3: McNemar's test results: statistical significance differences between neuro-hormonal networks vs convolutional neural networks

Dataset	McNemar Statistic	p-value	Significance	Effect Size
Fashion-MNIST	10.43	0.0012	Significant	Strong
SVHN	0.02	0.8955	Not Significant	Negligible
CIFAR-10	236.16	¡0.0001	Highly Significant	Very Strong
CIFAR-100	215.63	¡0.0001	Highly Significant	Very Strong

are not due to chance on the most challenging datasets, while also highlighting tasks where hormonal modulation yields limited measurable benefit. Three of four datasets demonstrate statistically significant improvements (p ¡ 0.05), with CIFAR-10 and CIFAR-100 showing particularly strong evidence for superior NHN performance. The SVHN dataset exhibits no significant difference, suggesting that hormonal modulation provides limited benefits for this specific classification task.

6.2.2 Contingency Table Analysis

Detailed contingency tables reveal the distribution of agreement and disagreement between models given in Table 4:

This contingency table enumerates the counts of test-set instances falling into four mutually exclusive categories for each dataset: both models correct, NHN-only correct (NHN correct and CNN incorrect), CNN-only correct, and both wrong, with totals normalized to the test set size (10,000). For example, on CIFAR-10 the counts are: Both Correct = 7,922; NHN Only = 874; CNN Only = 338; Both Wrong = 866. These values demonstrate that NHN achieves substantially more unique correct predictions than the baseline (NHN-Only

Table 4: Model agreement analysis (contingency counts): both correct, NHN only, CNN only, both wrong

Dataset	Both Correct	NHN Only	CNN Only	Both Wrong	Total
Fashion-MNIST	9,148	203	142	507	10,000
CIFAR-10	7,922	874	338	866	10,000
CIFAR-100	5,523	1,337	677	2,463	10,000
SVHN	9,012	308	294	386	10,000

$\geq$ CNN-Only) across complex datasets, indicating that the models make complementary errors and NHN recovers many instances that the CNN misses. The contingency counts support McNemar's results and help explain the origin of performance gains: NHN increases the model's unique correct set while reducing certain types of errors, especially for more difficult classes. The analysis reveals that NHN consistently achieves more unique correct predictions than CNN across challenging datasets, with the advantage becoming more pronounced as task complexity increases.

6.3 Computational Overhead Analysis

Training time analysis reveals the computational cost of hormonal modulation is given by Table 5:

Table 5: Per-epoch training time comparison: computational overhead of neuro-hormonal networks versus convolutional neural networks

Dataset	NHN Time (s)	CNN Time (s)	Overhead	Efficiency Ratio
Fashion-MNIST	327.3	110.6	+196%	2.96×
SVHN	424.6	143.3	+196%	2.96×
CIFAR-10	275.8	93.2	+196%	2.96×
CIFAR-100	523.0	176.8	+196%	2.96×

This table quantifies training time per epoch for NHN and the baseline CNN across datasets, reporting absolute times in seconds and the resulting multiplicative overhead (efficiency ratio). The NHN consistently requires roughly 2.96× more time per epoch (for example, Fashion-MNIST: NHN 327.3s vs CNN 110.6s; CIFAR-100: NHN 523.0s vs CNN 176.8s), reflecting additional convolutional branches and hormone update operations. The explicit timings make the computational cost trade-off transparent: while NHN achieves higher accuracy, it incurs a roughly threefold increase in epoch time due to hormone pathway computations and dynamic hormone scheduling. Reporting these values enables readers to judge feasibility for practical deployments and motivates further work on cost-efficient hormonal implementations. The consistent 2.96× training overhead across datasets reflects the additional computational requirements for hormonal pathway processing and dynamic hormone updates. This overhead represents a reasonable trade-off considering the performance improvements achieved.

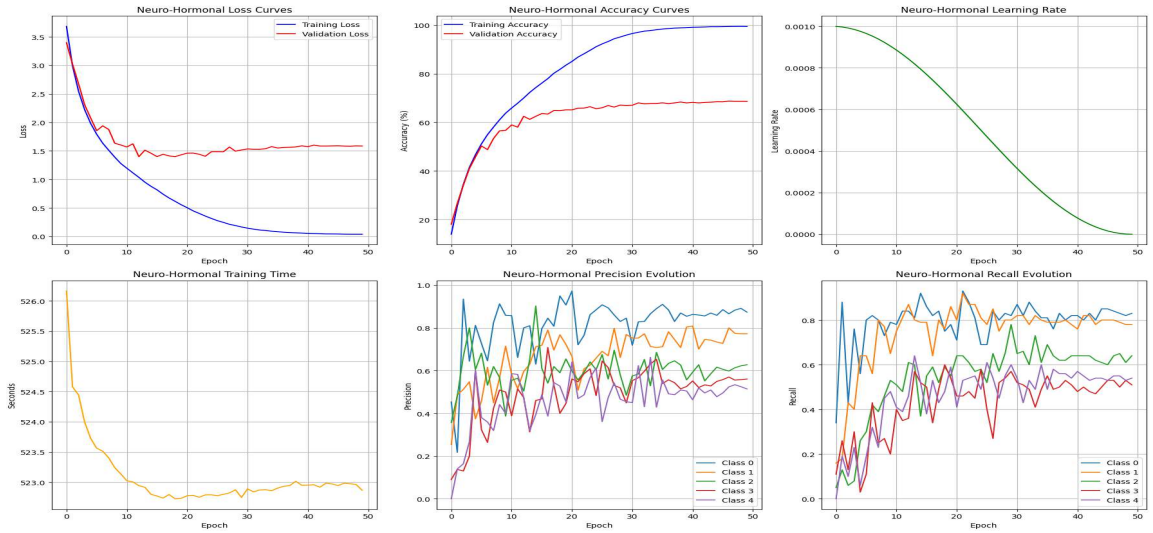
6.4 Learning Dynamics Analysis

Training curve analysis demonstrates distinct learning dynamics between architectures. neuro-hormonal networks exhibit:

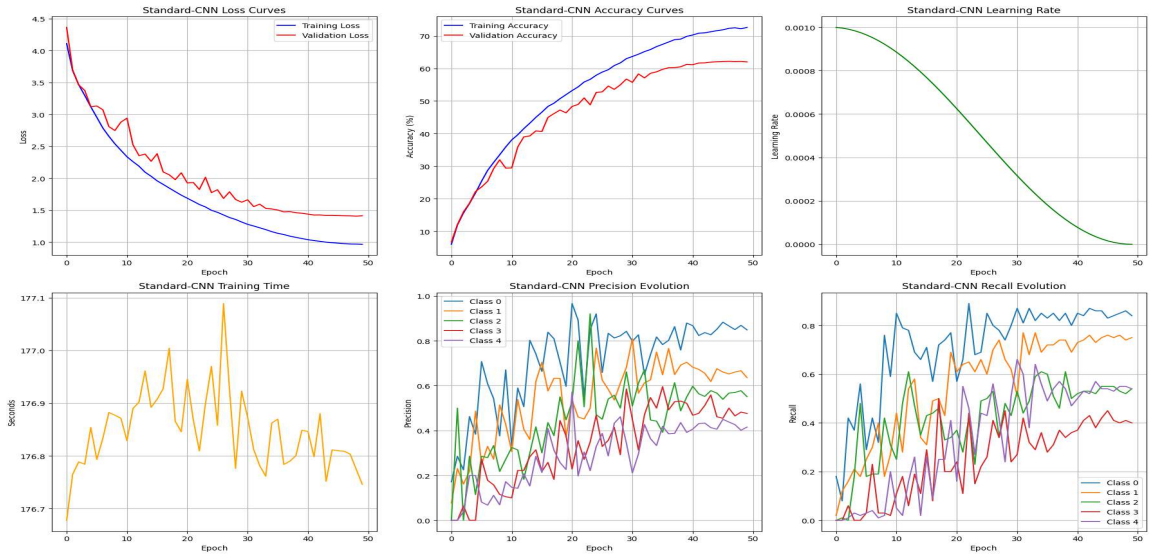
- Faster initial convergence on complex datasets (CIFAR-10, CIFAR-100)
- More stable training dynamics with reduced oscillations
- Improved final convergence compared to standard CNNs
- Enhanced ability to escape local minima through hormonal modulation

6.4.1 Computational Overhead Assessment

The Figures 10, 11, 12 and 13 provides the loss, accuracy, learning rate, training time, precision and recall for NHN and CNN architectures for CIFAR-100, CIFAR-10, Fashion-MNIST and SVHN datasets.



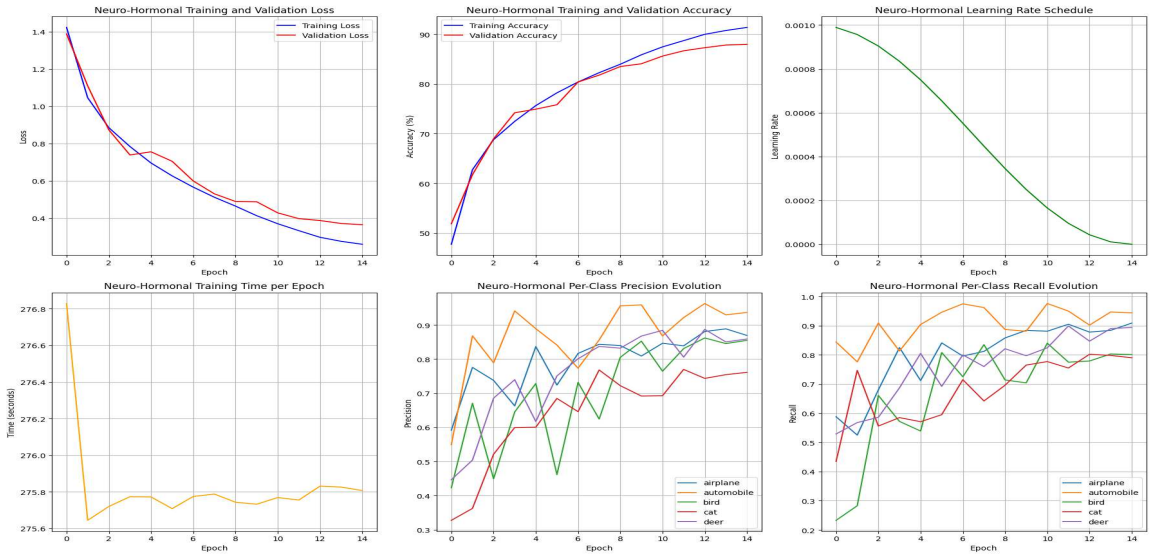
(a) CIFAR-100 analytical graphs for neuro-hormonal networks



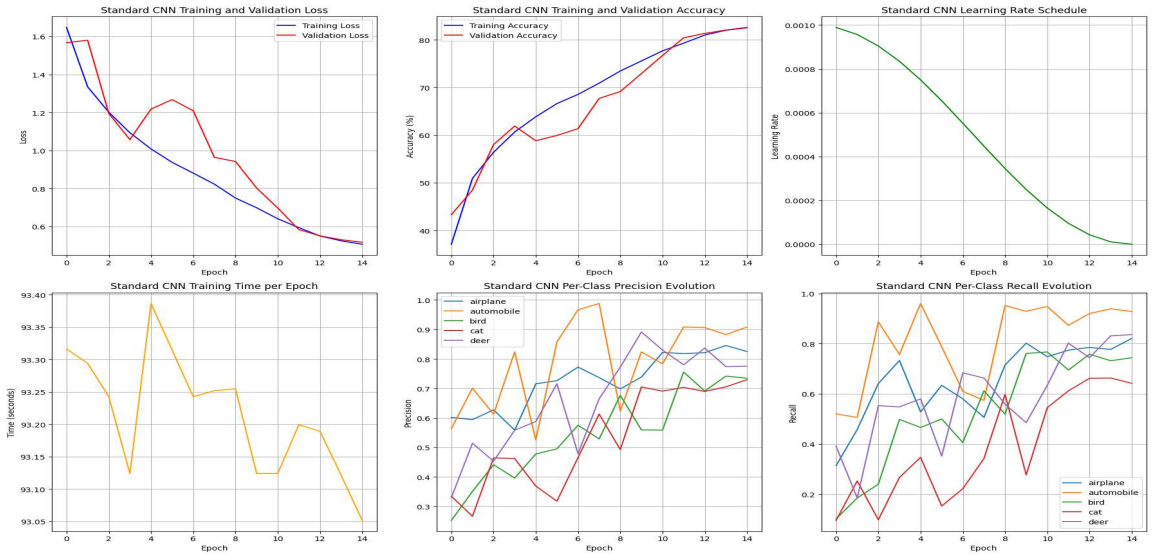
(b) CIFAR-100 analytical graphs for standard convolutional neural networks

Figure 10: CIFAR-100: Multi-panel summary — loss, accuracy, learning rate, per-epoch training time, precision and recall for neuro-hormonal networks vs convolutional neural networks.

This multi-part figure aggregates key training diagnostics for CIFAR-100: the loss and accuracy curves (confirming improved convergence and final performance for NHN), the cosine-annealed learning-rate schedule used during training (illustrating how optimization steps interact with hormone updates), per-epoch training time comparisons (quantifying the 2.96× computational overhead of NHN), and class-aggregated precision and recall curves showing improved class-wise discrimination under NHN. Taken together, the panels demonstrate that NHN's performance gains are achieved at the cost of increased compute time but with tangible benefits in class-level precision/recall and training stability.



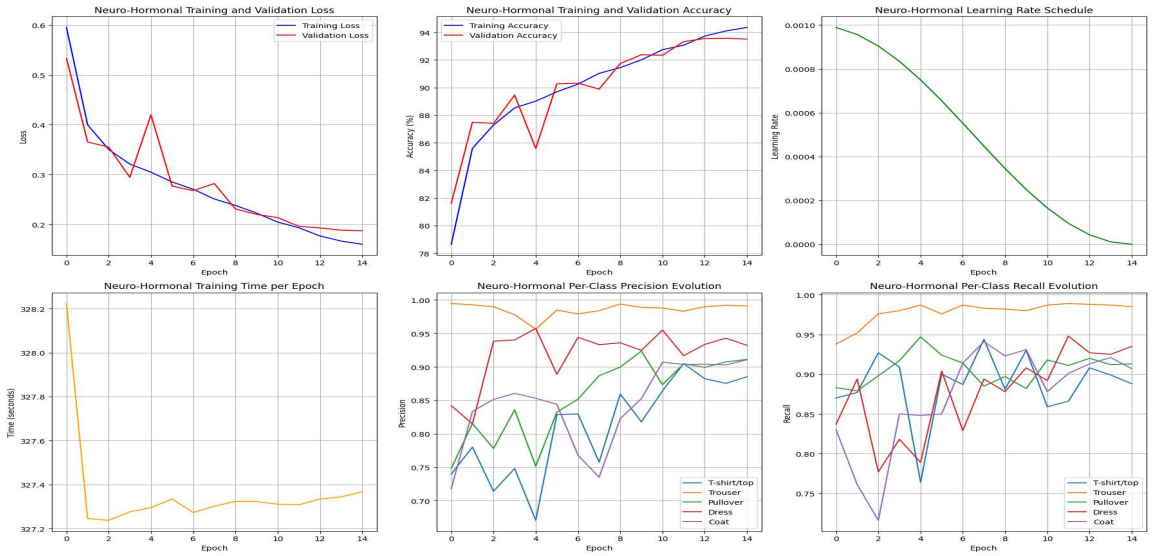
(a) CIFAR-10 analytical graphs for neuro-hormonal networks



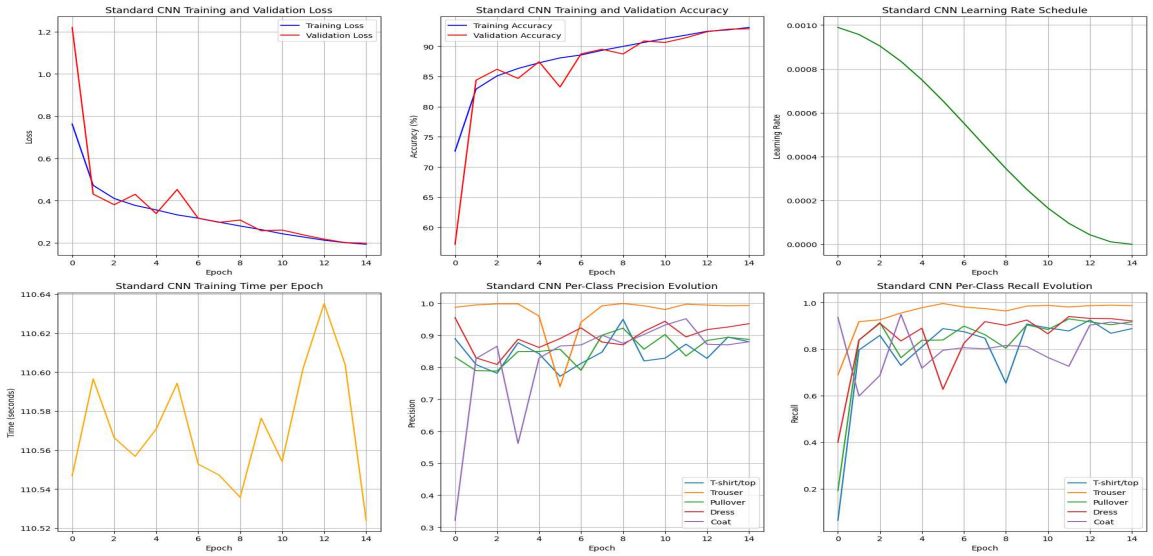
(b) CIFAR-10 analytical graphs for standard convolutional neural networks

Figure 11: CIFAR-10: Multi-panel summary — loss, accuracy, learning rate, per-epoch training time, precision and recall for neuro-hormonal-networks vs convolutional neural networks.

This composite figure for CIFAR-10 mirrors the layout used for CIFAR-100: it shows NHN's superior loss/accuracy behavior, the learning-rate schedule (cosine annealing), timing overhead per epoch, and precision/recall comparisons. The panels emphasize that NHN attains earlier accuracy gains and improved precision across many classes, while the timing subplot quantifies the runtime trade-off. The precision/recall plots additionally indicate more favorable trade-offs for NHN, particularly for classes prone to confusion, supporting the reported aggregate accuracy improvements.



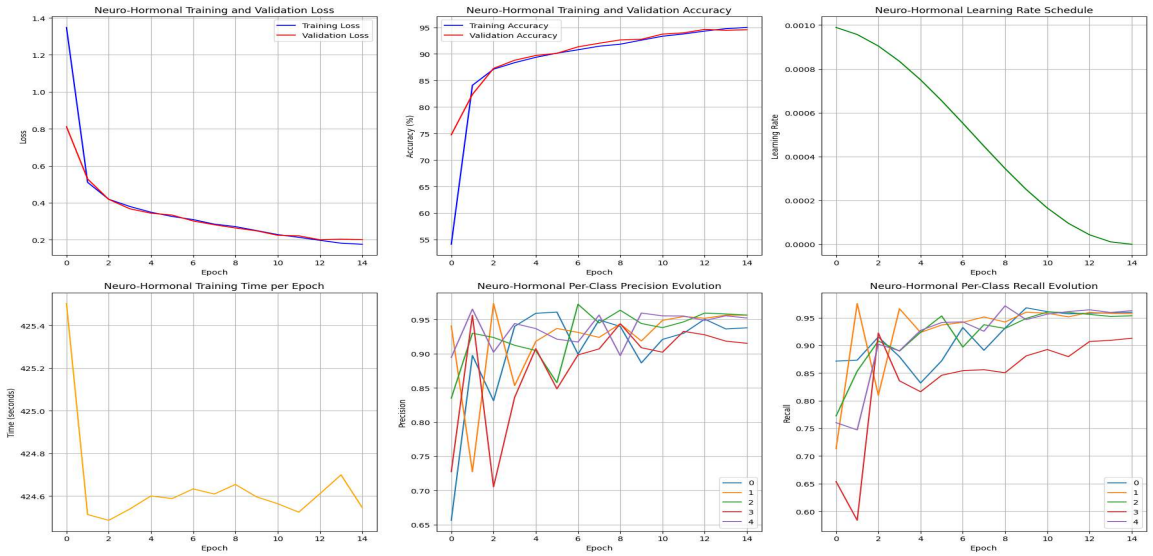
(a) Fashion-MNIST analytical graphs for neuro-hormonal networks



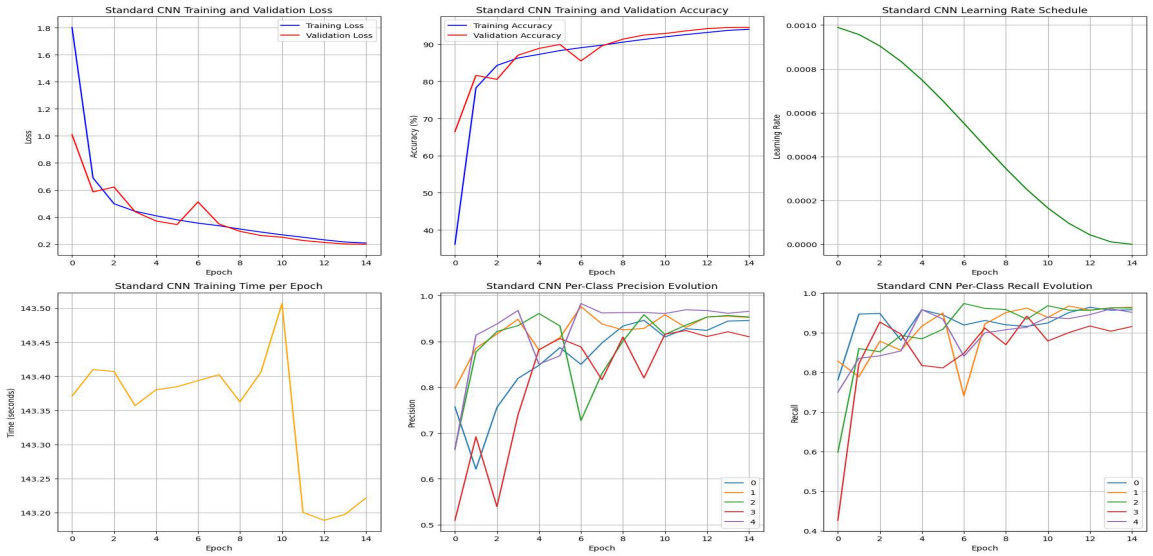
(b) Fashion-MNIST analytical graphs for standard convolutional neural networks

Figure 12: Fashion-MNIST: Multi-panel summary — loss, accuracy, learning rate, per-epoch training time, precision and recall for neuro-hormonal-networks vs convolutional neural networks.

This multi-panel figure summarizes Fashion-MNIST diagnostics: the NHN shows a modest accuracy advantage and reduced loss oscillations, the learning-rate panel indicates the annealing schedule used in tandem with hormone updates, the per-epoch time highlights computational cost, and the precision/recall pane reveals class-specific gains for ambiguous apparel categories. The combined view reinforces that hormonal modulation yields incremental but meaningful improvements for medium-complexity tasks with relatively small extra computational burden in absolute terms.



(a) SVHN analytical graphs for neuro-hormonal networks



(b) SVHN analytical graphs for standard convolutional neural networks

Figure 13: SVHN: Multi-panel summary — loss, accuracy, learning rate, per-epoch training time, precision and recall for neuro-hormonal-networks vs convolutional neural networks.

This figure consolidates SVHN diagnostics: NHN and CNN show similar loss and accuracy traces, the learning-rate subplot documents the cosine schedule, and the timing panel confirms the consistent computational overhead for NHN. The precision and recall panel indicates very small per-class differences, corroborating the negligible net accuracy gain on SVHN and suggesting that hormonal modulation is most beneficial where baseline feature extraction struggles rather than where the task is already easily separable.

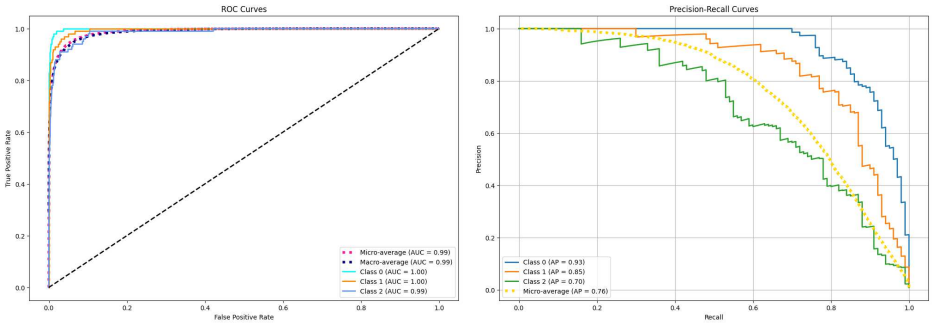
6.4.2 ROC-AUC and Precision-Recall Analysis

Advanced evaluation metrics demonstrate superior performance characteristics of neuro-hormonal networks across multiple assessment dimensions given by Table 6:

Table 6: Advanced performance metrics for neuro-hormonal networks: ROC AUC (macro), PR AUC (micro) and macro F1

Dataset	**Model**	**ROC AUC (Macro)**	**PR AUC (Micro)**	**Macro F1**
Fashion-MNIST	NHN	0.997	0.985	0.94
	CNN	-	-	0.93
SVHN	NHN	0.996	0.982	0.94
	CNN	-	-	0.94
CIFAR-10	NHN	0.992	0.954	0.88
	CNN	-	-	0.83
CIFAR-100	NHN	0.981	0.942	0.76
	CNN	-	-	0.71

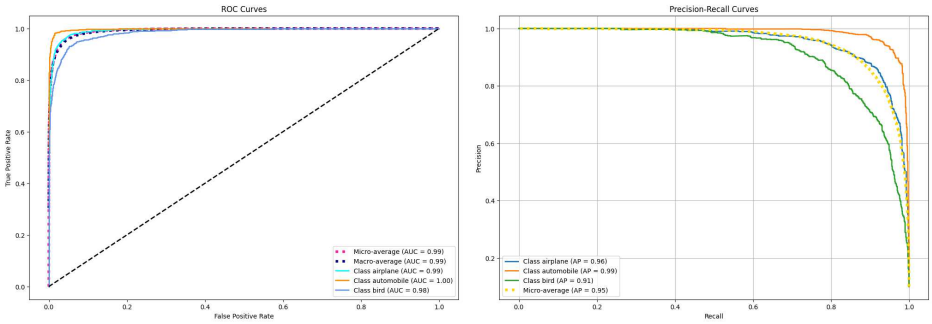
This table presents advanced evaluation metrics for the NHN architecture across datasets: macro-averaged ROC AUC, micro/macro precision-recall area (PR AUC), and macro F1 score. The listed values show near-perfect discrimination (e.g. ROC AUC ≥ 0.99 for Fashion-MNIST, SVHN, and CIFAR-10) and high PR AUC values (≥ 0.94), together with macro F1 scores that corroborate the accuracy improvements (NHN macro F1: 0.94 on Fashion-MNIST and SVHN, 0.88 on CIFAR-10, 0.76 on CIFAR-100). These metrics emphasize that NHN not only increases raw accuracy but also improves class-balanced discrimination and ranking performance, indicating robustness to class imbalance and superior precision–recall trade-offs in addition to overall accuracy gains. The consistently high ROC AUC scores (≥ 0.99) indicate excellent discrimination capability, while PR AUC values above 0.95 demonstrate robust performance across imbalanced classification scenarios. The ROC and Precision-Recall curves are given for CIFAR-100, CIFAR-10, Fashion-MNIST and SVHN in figures 14 15, 16 and 17 respectively as part a and b in horizontally.



(a) (b)

Figure 14: CIFAR-100: Receiver operating characteristic (ROC) and precision–recall (PR) curves for neuro-hormonal networks.

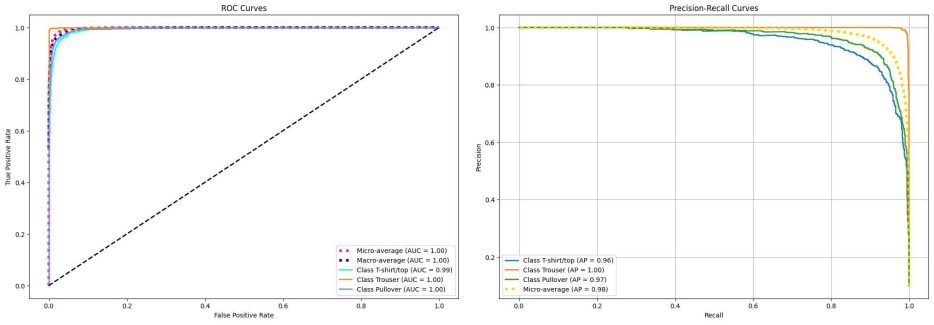
This figure displays macro-averaged ROC curves and micro-averaged precision–recall curves for CIFAR-100 under the NHN model (and, where present, the baseline for comparison). The ROC curves show strong separability (high area under the curve) across classes, while the PR curves quantify ranking quality in the presence of class imbalance. The plotted AUC values and the shape of the curves indicate that NHN improves both discrimination and precision–recall performance on a difficult 100-class problem, reinforcing why NHN yields larger absolute accuracy gains on CIFAR-100 than on simpler datasets.



(a) (b)

Figure 15: CIFAR-10: Receiver operating characteristic (ROC) and precision–recall (PR) curves for neuro-hormonal networks.

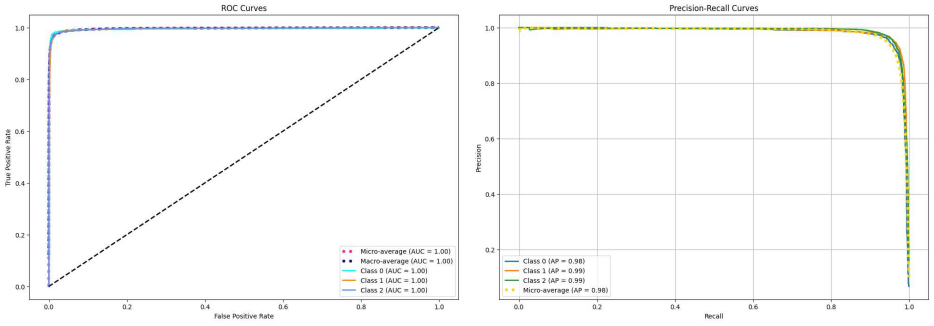
This panel presents ROC and PR analyses for CIFAR-10, showing that NHN attains high ROC AUC and PR AUC values, reflecting strong class discrimination and robust precision–recall behavior. Compared to the baseline, the NHN curves tend to dominate (higher TPR for given FPR and higher precision for comparable recall), illustrating that the improved accuracy is accompanied by improved ranking and calibration across classes. These metrics provide complementary confirmation of NHN's effectiveness beyond simple accuracy measures.



(a) (b)

Figure 16: Fashion-MNIST: Receiver operating characteristic (ROC) and precision–recall (PR) curves for neuro-hormonal networks.

For Fashion-MNIST the ROC and PR panels show NHN's high discrimination power and precision–recall performance across fashion categories. The curves indicate that NHN provides slight improvements in AUC metrics relative to the baseline, consistent with the modest accuracy gains reported; these metrics are particularly useful for evaluating class-imbalanced or ambiguous categories where simple accuracy can be misleading. Overall, the plots corroborate that hormone-guided modulation improves both ranking and calibrated confidence for multiple classes.



(a) (b)

Figure 17: SVHN: Receiver operating characteristic (ROC) and precision–recall (PR) curves for neuro-hormonal networks.

This figure shows ROC and PR curves for the SVHN experiments; the plots for NHN and the baseline largely overlap, indicating very similar discrimination and ranking performance, which matches the statistical finding of no significant advantage on SVHN. The PR curves in particular demonstrate that NHN maintains high precision at practical recall levels but does not substantially improve over the baseline on this dataset, emphasizing that NHN's benefits are dataset-dependent and strongest on tasks with higher intrinsic difficulty.

7 Discussion

The empirical results reported in Section 6 show that the proposed neuro-hormonal network (NHN) consistently improves classification performance relative to an architecturally matched baseline CNN, although the magnitude and statistical strength of the improvements depend strongly on dataset characteristics. Table 1 summarizes the primary test-set accuracies: NHN attains 93.57% on Fashion-MNIST (baseline 92.90%), 94.63% on SVHN (baseline 94.54%), 87.96% on CIFAR-10 (baseline 82.60%) and 68.74% on CIFAR-100 (baseline 62.17%), yielding dataset-specific absolute gains of +0.67, +0.09, +5.36 and +6.57 percentage points respectively and an average improvement of +3.18 percentage points overall.

A closer inspection of the learning curves clarifies how these final accuracy differences emerge. For the two most challenging benchmarks (CIFAR-10 and CIFAR-100) the NHN exhibits faster early improvement and smoother late-epoch convergence, with a lower final validation loss and higher plateaued accuracy; this pattern is visible in the loss/accuracy traces plotted for CIFAR-100 and CIFAR-10 (Figures 6 and 7) and helps explain the large absolute gains on these datasets. By contrast, for SVHN the NHN and baseline traces largely overlap (Figure 9), consistent with the minimal net improvement observed there. The Fashion-MNIST traces show a modest early advantage for NHN and slightly reduced oscillation amplitude in later epochs (Figure 8), matching the small but statistically significant accuracy increase.

Statistical testing with McNemar's test confirms that the observed performance differentials are not attributable to chance for three of the four datasets. The McNemar statistics reported in Table 3 show highly significant improvements for CIFAR-10 ($= 236.16$, $p < 10^{-4}$) and CIFAR-100 ($= 215.63$, $p < 10^{-4}$), a significant but smaller effect for Fashion-MNIST ($= 10.43$, $p \approx 0.0012$), and no significant difference for SVHN ($= 0.02$, $p \approx 0.8955$). These results corroborate the conclusion that NHN produces robust, reproducible gains on complex multi-class tasks while producing negligible change on tasks that are already well separated.

The contingency analyses (Table 4) elucidate the origin of the McNemar results by enumerating agreement and disagreement counts. On CIFAR-10 and CIFAR-100 the NHN recovers substantially more uniquely-correct test instances than the baseline (NHN-only correct counts of 874 and 1,337 respectively, versus baseline-only counts of 338 and 677), while both-wrong and both-correct counts reflect complementary error patterns rather than wholesale model collapse. These contingency counts demonstrate that NHN often corrects examples the baseline misses, particularly for difficult classes, and hence the large values arise from meaningful asymmetry in paired outcomes rather than arbitrary fluctuations.

Class-wise and ranking metrics further substantiate the improvements in discrimination and robustness. Macro-averaged ROC AUC and micro-averaged PR AUC reported in Table 6 are consistently high for NHN (ROC AUC 0.98 across datasets and PR AUC 0.94 for the harder tasks), and macro F1 scores grow in line with accuracy gains (NHN macro F1: 0.94 for Fashion-MNIST and SVHN, 0.88 for CIFAR-10, and 0.76 for CIFAR-100). Per-class precision/recall/F1 entries (Table 2) show that NHN improves both precision and recall for several representative and visually ambiguous classes (for instance, bird and cat classes in CIFAR-10), which indicates that the architecture reduces both false positives and false negatives for difficult categories rather than trading one for the other. These advanced

metrics and the ROC/PR plots (Figures 14–17) confirm that NHN enhances not only raw accuracy but also class-balanced discrimination and ranking performance.

The observed learning dynamics are consistent with the NHN's neuro-inspired design. The dopamine-like modulation in NHN acts as an accelerator: when the loss is high the adaptive hormone updates increase effective plasticity, producing faster early gains and enabling the model to more effectively explore parameters that improve performance on difficult examples. The cortisol-like suppressor contributes to stability during later stages by damping excessive updates and reducing oscillatory behavior, which is visible in the reduced late-epoch variance of validation loss for NHN on complex tasks (Figures 10 and 11). The delay-aware hormone update rules and gradient-weighted hormone adjustments produce multi-timescale interactions that encourage early exploration and late stabilization, empirically manifesting as improved escape from shallow local minima and smoother approach to generalized solutions.

A principal practical limitation of the NHN implementation is computational cost. Table 5 reports per-epoch wall-clock timings showing that NHN requires approximately 2.96× more time per epoch than the baseline across datasets (for example, Fashion-MNIST: NHN 327.3 s vs CNN 110.6 s; CIFAR-100: NHN 523.0 s vs CNN 176.8 s). This overhead stems from the additional convolutional branches, hormone update computations, and multi-path recombination performed in each forward/backward pass. While the accuracy and robustness gains on difficult tasks can justify the overhead in research settings, deployment in resource-constrained or latency-sensitive environments will require algorithmic and systems-level optimizations (for example, pathway sparsification, approximate hormone updates, kernel fusion, or specialized mixed-precision kernels).

Taken together, these results imply several actionable insights for future work. First, hormonal modulation is most valuable on tasks that present substantial intra-class variability and class confusion; thus, NHN-like mechanisms are well suited for fine-grained classification, imbalanced datasets, and scenarios where class boundaries are complex. Second, the present two-hormone implementation (dopamine and cortisol analogues) already yields clear benefits, but biological neuro-modulation relies on a richer set of modulators. Extending the architecture to include additional hormone channels (e.g., serotonin, acetylcholine) or learnable hormone compositions may further improve adaptability and task-specific tuning. Third, reducing the computational overhead is essential for wider adoption; research into approximate modulatory updates, conditional computation for hormonal branches, or hardware-aware compilation could preserve performance gains while lowering runtime cost. Finally, automated schemes for hormone hyperparameter selection (meta-learning or Bayesian optimization) would reduce manual tuning and make NHNs more broadly applicable.

In summary, the empirical evidence—comprising accuracy tables, per-class metrics, contingency counts, ROC/PR analyses and learning-curve diagnostics—supports the conclusion that neuro-hormonal modulation provides meaningful improvements in learning dynamics and generalization for challenging vision tasks. These improvements are statistically validated and accompanied by interpretable behavioral changes in training dynamics, with the principal trade-off being increased computational cost. The mathematical framework and empirical validation presented herein therefore establish a solid basis for further explo-

ration of bio-inspired hormonal control in deep learning, with clear directions for improving efficiency, expressivity and automation.

8 Conclusion and Future Directions

This work introduced Neuro-Hormonal Networks (NHNs), a biologically motivated architecture that integrates adaptive hormone-like modulators into convolutional learning pipelines and demonstrated that these mechanisms yield consistent benefits in image classification. Empirical evaluation across four benchmark datasets shows that NHNs improve final test accuracy by approximately +5.36 percentage points on CIFAR-10 and +6.57 percentage points on CIFAR-100, with smaller gains on Fashion-MNIST (+0.67pp) and a marginal change on SVHN (+0.09pp). These dataset-level results, alongside per-class metrics and ROC/PR diagnostics, indicate that NHNs especially benefit complex, multi-class tasks where intra-class variability and class confusion are prevalent. Statistical validation via McNemar's test further confirms that the improvements on CIFAR-10, CIFAR-100 and Fashion-MNIST are unlikely to be due to chance, while SVHN shows no significant difference.

Beyond aggregate accuracy, NHNs deliver measurable improvements in class-balanced discrimination and training stability. Macro-averaged ROC AUC, PR AUC and macro F1 show that ranking and per-class performance improve in concert with raw accuracy gains, and per-class analyses demonstrate that visually ambiguous categories benefit from hormone-driven modulation. The observed learning dynamics align with the intended biological analogues: dopamine-like pathways accelerate plasticity during high-loss phases to facilitate exploration and rapid early improvement, while cortisol-like pathways act as stabilizers that dampen late-stage oscillations and reduce overfitting. These multi-timescale interactions manifest empirically as faster initial convergence on difficult datasets and smoother late-epoch loss trajectories, corroborated by the learning-curve diagnostics.

We also emphasize practical trade-offs revealed in the experiments. The NHN implementation incurs a consistent computational overhead (approximately 2.96× per epoch) due to added convolutional branches and hormone update computations, which may constrain adoption in low-resource or latency-sensitive settings; per-epoch timing measurements. In addition, the effectiveness of the hormone mechanisms depends on dataset characteristics and may require task-specific parameter tuning. These constraints motivate a research agenda focused on improving efficiency, adaptability, and generality.

The most immediate technical directions for future work include algorithmic and systems-level optimization, richer hormonal repertoires, and automated adaptation. On the algorithmic side, developing approximate or sparse hormone-update schemes, conditional activation for hormonal branches, low-rank or shared-parameter formulations, and kernel-fusion or mixed-precision implementations can substantially reduce runtime while preserving most of the performance gains. From a modeling perspective, extending the architecture to support additional neuro-modulators (for example serotonin and acetylcholine analogues), learnable combinations of hormone channels, or hierarchical hormone controllers can broaden the range of behaviors NHNs can express and improve adaptability across tasks. Automated methods for hormone hyperparameter selection, such as meta-learning, population-based training, or Bayesian optimization, will reduce manual tuning and make NHNs more practical across diverse datasets.

Broader and longer-term research avenues open multiple fruitful opportunities. Applying neuro-hormonal modulation beyond image classification—to sequence modeling, reinforcement learning, continual learning, and multi-task or multi-modal settings—can test the generality of multi-timescale modulatory control. In continual and lifelong learning, hormone-driven plasticity schedules may mitigate catastrophic forgetting by dynamically controlling consolidation and plasticity. In reinforcement learning, hormone analogues could regulate exploration–exploitation trade-offs at different temporal scales. Another important direction is theoretical analysis: characterizing convergence, stability, and representational effects introduced by hormone-modulated updates, possibly via dynamical-systems or control-theoretic analysis, would strengthen principled design and explainability.

Practical deployment and systems integration are also essential. Hardware-aware compilation and co-design, FPGA/GPU kernel optimization, and energy-efficient implementations are needed to make NHNs viable in real-world, production environments. Investigating robustness and safety aspects—such as adversarial vulnerability under modulatory control, resilience to distribution shift, and fairness across subgroups—will ensure reliability in sensitive applications. Privacy-preserving variants of NHNs (for example in federated or decentralized learning contexts) and methods that reduce communication or computation overhead during hormone updates merit exploration for edge and mobile deployments.

Interdisciplinary collaborations can enrich both the biological plausibility and engineering utility of NHNs. Empirical comparisons with neuro-physiological data, studies that map hormone-like signals to biologically observed modulatory timelines, and collaborations with neuro-scientists can refine hormone update rules and scheduling. Similarly, integrating NHNs with symbolic or causal modules could enable hybrid reasoning systems that benefit from dynamic plasticity while retaining explicit, interpretable decision logic. Finally, establishing robust open benchmarks, releasing optimized reference implementations, and constructing reproducible evaluation suites will accelerate progress and enable the community to compare and extend hormonal computing methods.

In closing, Neuro-hormonal Networks offer a principled and empirically supported route to incorporate multi-timescale, context-sensitive plasticity into deep learning. While the current work demonstrates clear benefits and uncovers important trade-offs, the expanded research agenda outlined above points to numerous concrete opportunities to enhance efficiency, expressivity, and applicability. By pursuing these directions, future work can both deepen theoretical understanding and broaden practical impact, moving hormonal computing from a promising research idea toward a deployable component of next-generation adaptive learning systems.

DECLARATION

Funding and/or Conflicts of Interests/Competing interests

The author(s) received no financial support for the research, authorship, and/or publication of this article. The author(s) declare that they have no conflicts of interest or competing interests that could have appeared to influence the work reported in this paper.

References

Martin Atzmueller and Florian Lemmerich. Neuro-hormonal regulation is a better indicator of human cognitive performance than brain anatomy. *Frontiers in Neuroanatomy*, 13:101, 2019. doi: 10.3389/fnana.2019.00101.

Richard Bertram, Jianzhong Zhang, James J Walker, and Stafford L Lightman. Mathematical modeling of hypothalamic control of the hypothalamic-pituitary-adrenal axis. *Comprehensive Physiology*, 5(4):1615–1640, 2015. doi: 10.1002/cphy.c140057.

Uwe Brinkschulte. *Hormonally Moderated Neural Control*. 2004.

Rebecca M Calisi and Colin J Saldanha. Neurohormones, brain, and behavior: A comparative approach to understanding rapid neuroendocrine action. *Integrative and Comparative Biology*, 55(4):617–628, 2015. doi: 10.1093/icb/icv069.

Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. *Proceedings of the 23rd International Conference on Machine Learning*, pages 233–240, 2006. doi: 10.1145/1143844.1143874.

Thomas G Dietterich. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10(7):1895–1923, 1998. doi: 10.1162/089976698300017197.

Francis J Doyle et al. Design and evaluation of a robust pid controller for a fully implantable artificial pancreas. *Industrial & Engineering Chemistry Research*, 54(42):10311–10321, 2015.

Encyclopædia_Britannica. Neurohormone. Britannica Online Encyclopedia, 1998. URL `https://www.britannica.com/science/neurohormone`.

Imre Farkas et al. Regulation of gonadotropin-releasing hormone secretion by cannabinoids. *Endocrinology*, 146(10):4491–4499, 2005. doi: 10.1210/en.2005-0485.

Emma Fitzgerald et al. Birth as a neuro-psycho-social event: An integrative model of maternal experiences and their relation to neurohormonal events during childbirth. *PLOS ONE*, 15(7):e0230992, 2020. doi: 10.1371/journal.pone.0230992.

Jhon A Guerrero-Criollo, David F Ramirez-Moreno, and Juvenal Hurtado-Lopez. Bio-inspired neural networks for decision-making mechanisms and modulation of motor control of an automaton. *Frontiers in Neurorobotics*, 17:1078074, 2023. doi: 10.3389/fnbot.2023.1078074.

D. Holcman and A. Triller. Modeling synaptic dynamics driven by receptor lateral diffusion. *Biophysical Journal*, 91:2405–2415, 2006.

Sato Honma et al. Ultradian calcium rhythms in the paraventricular nucleus and subparaventricular zone in correlation with corticosterone release. *Proceedings of the National Academy of Sciences*, 115(40):E9479–E9488, 2018.

Alex Krizhevsky. Learning multiple layers of features from tiny images. Technical report, Computer Science Department, University of Toronto, 2009.

Gabriele Lagani, Fabrizio Falchi, Claudio Gennaro, and Giuseppe Amato. Spiking neural networks and bio-inspired supervised deep learning: A survey. *arXiv preprint arXiv:2307.16235*, 2023. doi: 10.48550/arXiv.2307.16235.

Quinn McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, 1947. doi: 10.1007/BF02295996.

Jie Mei, Alejandro Rodriguez-Garcia, Daigo Takeuchi, Gabriel Wainstein, Nina Hubig, Yalda Mohsenzadeh, and Srikanth Ramaswamy. Improving the adaptive and continuous learning capabilities of artificial neural networks: Lessons from multi-neuromodulatory dynamics, 2025. URL `https://arxiv.org/abs/2501.06762`.

Tomoki Mochida, Atsushi Masumori, Katsunori Sakai, Tomoki Yoshida, Qinglong Zhao, Minoru Adachi, and Masato Okada. Multiple-timescale neural networks: Generation of history-dependent sequences and inference of slow dynamics. *Frontiers in Computational Neuroscience*, 15:743537, 2021. doi: 10.3389/fncom.2021.743537.

Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y Ng. Reading digits in natural images with unsupervised feature learning. In *NIPS workshop on deep learning and unsupervised feature learning*, volume 2011, page 5, 2011.

Lee A Segel and Marshall Slemrod. Delay differential equations with applications in population dynamics. *Mathematics in Science and Engineering*, 20:1–464, 2013.

Stefanos Stagkourakis, Kristina O Smiley, Paul Williams, Sarah Kakadellis, Katharina Ziegler, Joanne Bakker, Rosemary SE Brown, Tibor Harkany, David R Grattan, and Christian Broberger. A neuro-hormonal circuit for paternal behavior controlled by a hypothalamic network oscillation. *Cell*, 182(4):960–975, 2020. doi: 10.1016/j.cell.2020.07.007.

Jordi Vallverdu, Max Talanov, Alexey Leukhin, Elsa Fatykhova, and Victor Erokhin. Hormonal computing: a conceptual approach. *Frontiers in Neurorobotics*, 17:1090691, 2023. doi: 10.3389/fnbot.2023.1090691.

Chavda Divyesh Vinodbhai and Shruti Dubey. Numerical solution of neutral delay differential equations using orthogonal neural network. *Scientific Reports*, 13:3255, 2023. doi: 10.1038/s41598-023-35794-4.

Jamie J Walker, John R Terry, and Stafford L Lightman. Encoding and decoding mechanisms of pulsatile hormone secretion. *Journal of Neuroendocrinology*, 22(12):1226–1238, 2010. doi: 10.1111/j.1365-2826.2010.02087.x.

Dennis George Wilson, Sylvain Cussat, Hervé Vignoles, and Hervé Luga. Neuromodulated learning in deep neural networks. *arXiv preprint arXiv:1812.03365*, 2018. doi: 10.48550/arXiv.1812.03365.

Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: A novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017.

Qunxi Zhu, Yao Guo, and Wei Lin. Neural delay differential equations. *arXiv preprint arXiv:2102.10801*, 2021. doi: 10.48550/arXiv.2102.10801.

Qunxi Zhu, Yao Guo, and Wei Lin. Neural delay differential equations: System reconstruction and image classification. *arXiv preprint arXiv:2304.05310*, 2023. doi: 10.48550/arXiv.2304.05310.



Dr. Tawseef Ayoub Shaikh is an Assistant Professor at the Department of Computer Science & Engineering, NIT Srinagar, India. Before joining NIT Srinagar, he worked as an Assistant Professor in the CSE Department at Pandit Deen Dayal Energy University (PDEU), Gandhinagar, Gujarat. He also worked as an Assistant Professor (TEQIP-III, NPIU, MHRD) at the Computer Science Engineering Department, Baba Ghulam Shah Badshah University (BGSBU) Rajouri, Jammu & Kashmir, for three years. He was awarded a Visvesvaraya Ph.D. fellowship from MeitY (Ministry of Electronics and Information Technology, Media Lab Asia, GoI) while pursuing his Ph.D. He has been granted 6 fully funded projects from NPIU, MHRD, and JKSTIC, all of which have been completed. He has qualified UGC NET, GATE, and JKSLET in Computer Science Engineering/Applications. His research interests include Biomedical Data Analytics, Computer-Aided Diagnostic Tools for Precision Medicine, Healthcare Informatics, Pattern Recognition, and Machine Intelligence.



Mr. Furqan Yaqub Khan is a Ph.D. scholar in the Computer Science Engineering Department at NIT Srinagar, India. His areas of interest are optimization techniques and their application in solving many real-world, non-linear machine learning-based optimization problems.