**Science, Engineering and Technology** Vol. 5, No. 1, pp. 106-115
www.setjournal.com *https://doi.org/10.54327/set2025/v5.i1.245*

# A Computational Model of Attention-Guided Visual Learning in a High-Performance Computing Software System

Alice Ahmed[1], Md. Tanim Hossain[2]

[1]*Department of Software Engineering, Daffodil International University, Dhaka, Bangladesh.*
[2]*Information and Communication Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China.*

## Abstract

This research investigates transformer architectures in high-performance computing (HPC) software systems for attention-guided visual learning (AGVL). The study focuses on the effects of environmental factors and non-contextual stimuli on cognitive control. It reveals how attention increases responses to attentive stimuli, thereby normalizing activity across the population. Transformer blocks use parallelism and less localized attention than current or convolutional models. The study investigates the use of transformer topologies to enhance language modeling, focusing on attention-guided learning and attention-modulated Hebbian plasticity. The model includes an all-attention layer with embedded input vectors, non-contextual vectors containing generic task-relevant information, and self-attentional and feedforward layers. The work employs relative two-dimensional positional encoding to address the challenge of encoding two-dimensional data such as photographs. The feature-similarity gain model proposes that attention multiplicatively strengthens neuronal responses based on how similar their feature tuning is to the attended input. The attention-guided learning approach rewards learning with neural attentional response gain, which the network modifies via gradient descent to achieve the projected objective outputs. The study discovered that supervised error backpropagation and the attention-modulated Hebbian rule outperformed the weight gain rule on MNIST; however, concentration differed.

**Keywords**: *Computational Model, Attention-Guided Visual Learning, High-Performance Computing, Reinforcement Learning, Computer Vision*

## 1. Introduction

Higher-performance computing (HPC) is a system where multiple computers, servers, or workstations pool their resources to perform specific tasks. These resources can be on-premises, in the cloud, or a combination of both. Nodes in a cluster are individual computers that perform specific tasks, such as computation, storage, and networking. High-performance computing (HPC) is revolutionizing businesses by processing massive volumes of data at rapid rates. The market is estimated to reach USD 86.36 billion by 2030, with a CAGR of 7.7% between 2024 and 2030. Major developments including artificial intelligence (AI) insertion, cloud-based services, quantum information processing, energy-saving technologies, exascale computation, improved archived data, and loose resource advancement. Such advances propel HPC forward, allowing advances in scholarly investigation and use in industry. Organizations may use high-performance computing its fullest potential towards an unparalleled level of processing speed and efficacy. HPCs are used in various fields such as research, design, simulation, and BI. They are also crucial for societal functions like validation of credit card transactions, vehicle design evaluation, and weather prediction.

Corresponding author: *Alice Ahmed (aliceahmed.cn@gmail.com)*
Received: *15 October 2024*; Revised: *11 December 2024*; Accepted: *16 December 2024*; Published: *25 December 2024*

Despite having access to a wealth of new data, we must make choices and focus on the most important details. Attention is often seen as a selection process, but it is essential to make the most of the available resources and make informed decisions to maximize the benefits of HPC [1]. On top of that, our focus tends to gravitate toward details that are directly related to the current topic of discussion. This is due to our innate tendency to pay attention to details. There is a logic that elevates the importance of focus beyond its actual value. The use of neuronal populations enables the encoding of information pertinent to the present situation. It is indeed possible to achieve this. The use of populations of neurons is one mechanism that enables this. The visual cortex has been the primary focus of neurophysiologists' research since the turn of the twentieth century. This trend has persisted since the beginning of the 20th century [2]. They have mostly concentrated on this part of the person being investigated within their investigations. Acquiring a more thorough understanding of this process is their objective. Researching and understanding the method is one of several goals that have been considered for this study's execution. Although there has been a lot of research on visual attention, the processes that generate Attention-Guided Visual Learning (AVGL) remain a mystery. Figure 1 shows the architecture of AGVL. This remains the case despite a significant amount of research that primarily focuses on attention. Despite a significant amount of research on visual attention, this situation has been observed. Even though a significant amount of research has already been conducted, it is evident that this is the true nature of the issue. This is a result that has been noted. Using computer models designed to mimic the intricate connections seen in the brain might be an option for integrating attentional processes into the system. This may come to pass. One way to achieve this is by applying this technique. To provide a more concrete example, this might be achieved by focusing the network's efforts on data that is relevant to particular data sets [3]. I can assure you that this is within reach. These models, which have the potential to facilitate the implementation of attentional processes, make it feasible to achieve this. It is possible to achieve this. Indeed, we can achieve this. The execution of attentional processes is something that can be accomplished. On the other hand, existing attentional models either ignore attention altogether as a part of learning [4], misrepresent the visual cortex's biological learning processes [5], [6], [7], or don't address the question of how learning relates to known attentional modulations [8], [9]. Moreover, these models are finding application in a wide range of contexts.

High-performance computing (HPC) is a complex system that involves processing, networking, and storage for large, sophisticated projects. It is primarily used in clusters and distributed computing, which use shared computation to reduce latency [10]. Cloud-based HPC supports complex activities like data storage, networking, security, specialized computing resources, and AI applications with scalability and flexibility. It boosts R&D speed, performance, efficiency, cost savings, and fault tolerance [11]. HPC allows companies to quickly analyze data, generate new ideas, and make scientific discoveries. It is used in AI visualization, optimization, data analysis, prediction, and research, optimizing huge datasets, accelerating genomic sequencing, and making real-time predictions.

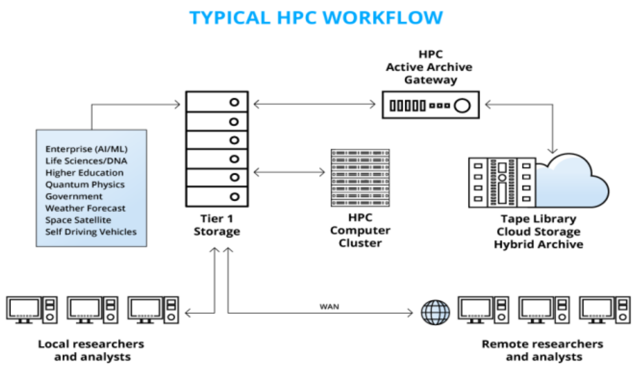


**Figure 1.** High Performance Computing (HPC) workflow architecture.

## 2. Background Study

Mathematical models support AEC building design, choices, investigation, and computational projections because powerful technology, especially HPCs, made PC professions easier to achieve. Because they offer a model representation of actual systems, computational models are crucial resources for comprehending and modifying systems. There are several uses for digital modeling, including solving challenging problems in reinforcement learning. The suggested training framework necessitates attention-guided learning and unusual information and makes use of intellectually modifying model-based reinforcement and theoretically appropriate components [12]. A computational model assessment requires output bias correction and a model representation of the real-world system. Using computational models to help students reflect and understand systems may be effective. One question provides a practical example of this paradigm in action [13]. Modify the model to highlight the issue. Remove components to see whether the model can achieve its aims. Draw a diagram or write down system connections. Computational models are evaluated by checking that each phase produces the intended result, comparing it to the real system, and making adjustments. Many computational thinking templates are available for middle school science classes. One student created a computer model to address illness transmission and social isolation [14]. AEC research uses computational exercises, mathematical modeling related to structural formulas, and ML to address complicated issues. Since this discovery, digital modeling has found numerous applications. Computer modeling is important because it mathematically simulates activities and physical actions. Since mankind began, science, technology, and business executives have utilized it, but its importance is growing [15]. AEC research uses computational exercises, mathematical modeling of structural equations, and ML to address complicated issues. Since its discovery, digital modeling has had numerous uses. Focus and firing variations in reinforcement learning are examined here. Trial and instruction, together with concentration activation regulation simulations, create a realistic training framework. Cognitive changes before incentives indicate neuronal plasticity, which drives the suggested training paradigm. The study lacks growth in systems, such as brain neurons; hence, linguistic proficiency ignores "accuracy." This study and artificial neural networks cannot detect whether brain transmitters that underpin relationships are repressive or joyous [16]. Cognitive vs. neural: the suggested training framework uses cognitive modulation, model-based reinforcement, and scientifically suitable components. This task requires attention-guided learning and uncommon information.

## 3. Method and Model

According to the top-down technique, information flows between regions of the cortex that are higher than usual, while information flows in the opposite manner from below; both strategies have an effect on cognitive control. Tasks requiring visual analysis to differentiate a range of attributes among several target possibilities require both types of information. Environmental features induce simultaneous top-to-bottom neural alternatives, while non-contextual "alarming" stimulation produces concentrated changes across different dorsal regions [17].

The top brain regions are used to judge object-specific concern, though the prefrontal zone evaluates learning. The sensory system, the parietal lobe, and the frontal lobe—particularly the inner ear—are evaluated for the capacity to concentrate. While spatial signaling directs the focus to a particular location throughout the area of purpose, features transmission contains targeted color, which includes movement signals. Goal signaling enhances accuracy and response time in a variety of visual-based activities. However, when the degree of threat goes up, cognitive gains diminish when working for difficult tasks such as low-contrast object detection. To concentrate on the correct visual field of view, as seen in Figure 2, perception tests used positional and characteristic fluctuations, cognitive fluctuations among signals and the aim, and more.

While some studies have demonstrated that people's attention increases even when there is significant environmental disturbance [9], other studies have found that the opposite occurs when there is no interruption [14]. The following offers more proof, showing that paying attention reduces the impact of extraneous signals therefore boosts the response elicited by focused stimulation. Signaling actions, as opposed to non-cueing duties, have been shown to improve both the ultimate result and actual objective contrast [6]. Whatever is occurring, it is a result of our perspective altering how we see our surroundings. Thus, versus giving a faithful democracy, that should focus on matters that we really think helpful or important. Information pertaining to the future is received by visual systems through its retinas, namely between the cerebellum and frontal lobe regions. Attention is simulated in this investigation. Data is sent to above GUI association regions by systems in the posterior and superior visual cortex. This feedforward movement is conveyed and increases the neuronal diversity produced by both feature-based and spatial concentration. The cortical area, central intraparietal domain, as well as frontal eye fields (FEF), and this are involved in psychological objectives and preparation, including decision-making, also affect curiosity [18]. The information created by the frontal area is conveyed to the hearing areas through the flow of vision or other channels. These pathways tell lower sensory regions about spatial and morphological biases. Our gazes are guided by visual spatial priority systems, which identify socially meaningful sites. Importance patterns may exist in both the frontal eye fields and the middle brain area, extending beyond the human visual function.

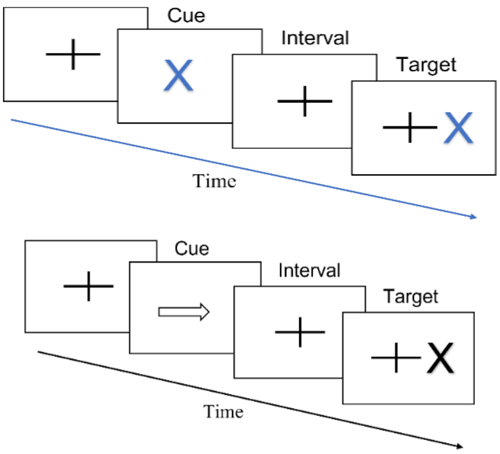


**Figure 2.** Cognitive fluctuations among signals proper visual field of view.

Each neural concentration fluctuation indicates the neuron's field of vision. Neural receptors extend every optical forwarding circulation as each place assimilates knowledge from the preceding dimension, similar to deep neural networks known as convolutional networks [19]. Furthermore, cognitive brain networks show differential activity in response to 'preferred' patterns. The capacity to decipher something is thought to improve with a more constrained adjustment range [20]. In contrast to non-attended jobs, visual stream neurons show higher firing rates, faster reaction times, and greater consistency after stimulus onset. Instruction amplifies neuronal community data, especially in demanding tasks, hence improving cognitive response through neural feed-forward flow and raising task complexity. Increased focus also reduces the latency between stimuli and brain responses, as demonstrated in Figure 3.

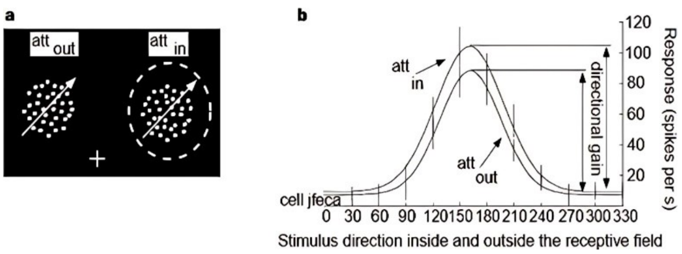


**Figure 3.** Orientation adjusting spectrum cognitive growth [17]: a) Flexible region or outdoors; b) Multi-scaled responses.

Experiments determine which differences in how we think about various aspects of things improve neural network performance, which aids in functions such as positional cognitive benefits. This is similar to spatial cognitive responses to stimuli in the receptive region as it reduces bilaterally difference links between neurons with similar tuning and improves rhythm synchronization [21], [22]. Neuronal responses tailored to the receiver field may be included in feature-based spatial attention. According to research, cognitive fluctuations vary, as detailed in the section below. Attentional response increases are most pronounced in regions that preferentially react to the target object, such as the fusiform facial area for attending faces. These non-space-dependent changes make firing patterns stronger when there are targets and weaker when there are distractions. This strengthens gamma oscillations.

These findings suggest that upper cortical object-based attention behaves similarly to feature-based attention. Many hypotheses are proposed to explain the experimental data, but none explain why some studies report contrast gains while others report contrast-tuning function response gains. There is no hypothesis that explains why a neuron's firing rate changes when it switches the focus between preferred and anti-preferred features [23], [24]. This model proposes that attention multiplies the response to the attending stimulus, but that neuronal activity is normalized across the population based on receptive field location and tuning [25]. The nonlinearity of the normalization model allows for fascinating discoveries of feed-forward flow attentional modulations, as shown in Figure 4. Microsimulation of lower cortical neurons demonstrated that changes in cross-area attention are caused by new connections to higher regions rather than changes in how neurons respond locally.

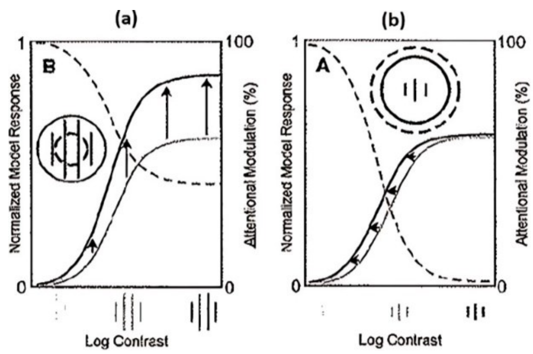


**Figure 4.** Visual focus modulates contrast tuning mechanisms [24]: a) Sensitivity improvement; b) Gaining comparison.

Some researchers propose a tuned normalization model, which balances everything cells' tuned responses to every stimulus with contrast cues in the receptive field. This model explains why different neurons provide varying attentional benefits. This approach uses attention (βA) to increase the rate of response and normalization intensity for attended stimuli [26], [27].

## 4. Attentional Mathematical Models

Based on the human visual system, convolutional networks interact with artificial neurons via weighted connections. Like the visual system, deep convolutional networks (DNNs) retain specialized data across large areas in successive layers. These networks simplify theoretical attentional hypothesis formulation and testing. However, previous computational research focused on object segmentation frameworks such as Mask R-CNN, which are computationally expensive, do not scale well with input size, and were not designed to mimic attentional processes for classification.

In supervised learning with differentiable loss functions for network weights, backpropagation can be used to teach soft attention computational models. As convolutional neural networks (CNNs) increase linearly with pixel count, intense attention was first used [28] to reduce the processing cost of large images. To identify many items, a more sophisticated variant continues this procedure until a projected stop-sign label is achieved. Hard and feature-based attention models in physiologically realistic computational systems differ. Based on prediction accuracy, feature-based attention models prioritize features. Weighting predictions from incoming text helps language models 'highlight' relevant words. Decoders in two recurrent networks apply state-dependent weighting to embedded words in the input vectors to predict.

Language tasks like translation and question-answering benefit from soft attentional models. Due to state linking, recurrent networks cannot train on extended sequences. Storing sequences in two dimensions may teach convolutional networks to "soft-attend" or give input sections greater weight. Due to CNN localization and constrained receptive fields, soft attention models cannot handle long-range dependent sequences. Transformer blocks employ parallelizability and less locally oriented attention than current or convolutional models. Weighted embedding word functions create key (K), value (V), and query (Q) matrices in self-attention blocks. All Attention Block outputs are collected, added to the sublayer via a residual link, normalized, and routed to the feedforward layer after each time step, as shown in Figure 5.

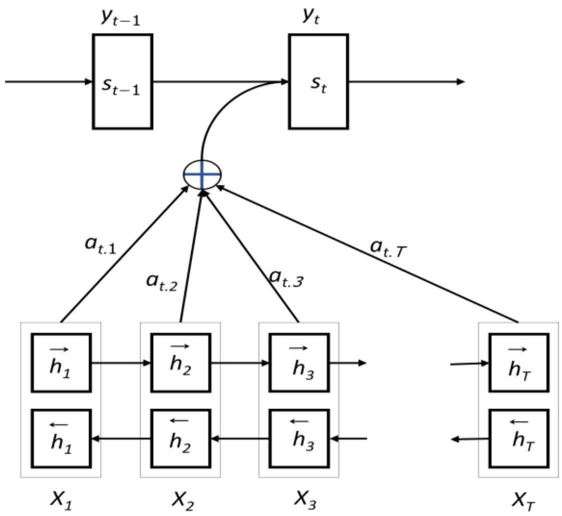


**Figure 5.** Two attention-implementing algorithms. Example of DRAM's rigorous concentration [1].

Many computational models have been constructed by changing or extending transformer architectures to improve language modeling. The model has an all-attention layer with embedded input vectors, non-contextual vectors with generic task-relevant information, and self-attentional and feedforward layers. Encoding two-dimensional data like photos is complicated and may not enhance performance. The current study uses relative two-dimensional positional encoding to overcome this problem. Positive findings from image classification datasets do not prove that self-attentional models mirror human attention. Retinal neurons cannot double their output because self-attention requires multiplication across variables.

The feature-similarity gain model says that attention multiplicatively makes neural responses stronger based on how similar their feature tuning is to the attended input. A recent study at the population level multiplicatively scaled the output of a pre-trained DNN feature map based on the category tuning value. The tuned attention approach enhanced image classification at all levels. Category-specific scaling of neuronal population activity enhances accuracy by identifying each category separately. The final layer's emphasis boosts attentional performance by

18.8% in integrated pictures and 22.8% in 2x2 image grids. Focusing on earlier stages reduced performance advantage. Tuning curves intensify as one moves through the brain's layers, indicating stronger tuning. Equation 1 tuned similarly and performed better when the average prediction-error gradient replaced feature map f for category c.

## 5. Attention-Guided Learning Method

Physiologically realistic visual system Attention-guided learning rewards learning with neuronal attentional response gain. To simulate attentional gain, the model multiplies neuron y activity by βy during training. An attentional gain in ReLU-activated neurons is influenced by prediction error δ feedback. This study optimized a network using reinforcement learning. During the action-selection phase, all β's begin at 0, indicating no initial attentional modulation. The network selects the top output unit with probability p and explores another unit with probability 1−p. From predicted unit s, the network calculates ϵs as a goal output. During the attentional phase, the network uses gradient descent to modify its gain to meet the predicted objective outputs. This allows attention to modify network activity before learning, much like the visual system does. Cross-entropy loss is a reward for mispredictions.

In the attentional phase, β is optimized to match target output ϵs for a certain number of iterations (t=0,...,T) called the attention span. The network's reward prediction error r σs affects whether learning enhances or inhibits attention-modulated network activity after obtaining a reward. One attention-guided learning hypothesis says postsynaptic attentional gain strengthens synapses, whereas the other says presynaptic activity and relative attention determine learning. The weight update ensures that neuron Y activates in the direction of its attentional gain in rewarded trials and reverses in unrewarded trials.

The learning rule Attention-modulated Hebbian plasticity provides pre- and post-synaptic feedback. Weight sign changes distinguish attention-guided learning from Equation (6)'s attentional term feedback. At a rate of σs, the hidden layer's feedback optimizes the attention-modulated learning rule to adjust attentional activity. Attention-modulated Hebbian plasticity Rule approximations that reduce computing costs may enhance categorization. This feedback rule optimizes attentional weight updates. Use the update Δβˆy = α by (δ) to estimate the neuron's attentional term βy. Using βˆy yin instead of βy in the forward pass results in neural output yout = g (1 + βˆyyin) · yin. The estimated feedback approach reduces computationally expensive division, as shown in Figure 6.

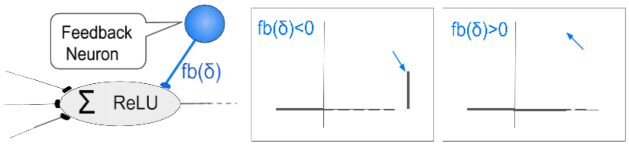


**Figure 6.** The impact of response fb(δ) affecting prediction errors δ for ReLU-activated neuronal cognitive development.

Due to diverse assumptions like enhanced attentional weight and attention-modulated Hebbian plasticity, attention-guided learning has two principles. The criteria were assessed using MNIST, which includes 70,000 greyscale images of handwritten digits at 28x28 pixels, and CIFAR10 and CIFAR100, which have 60,000 color photographs of objects and animals at 32x32x3 pixels. The CIFAR10 vehicle subset, CIFAR4, which includes the classes 'automobile,' 'airplane,' 'ship,' and 'truck,' was computationally tested using 10,000 pre-selected samples. Attention spans vary. T demonstrated how well many learning strategies were in categorizing. For a fair comparison, attention rates α were estimated as αˆ/T at the same global rate αˆ. A small neural network with three fully connected hidden layers triggered by ReLU and a linear output layer tested both learning algorithms on MNIST. Two 32x3-filter convolutional layers handled CIFAR4, 10, and 100 data.

Simulation results of the study are shown in Figure 7 and Figure 8, as well as in Table 1, Table 2, and Table 3. Gradient-weighted class activation mapping (GRAD-CAM) was used to illustrate network heat maps to test whether attention prioritizes significant data over irrelevant data. GRAD-CAM visualizes prediction-impacting regions using gradients from the final convolutional layer, which incorporates high-level spatial information. Class-discriminative visualization improves model behavior knowledge.

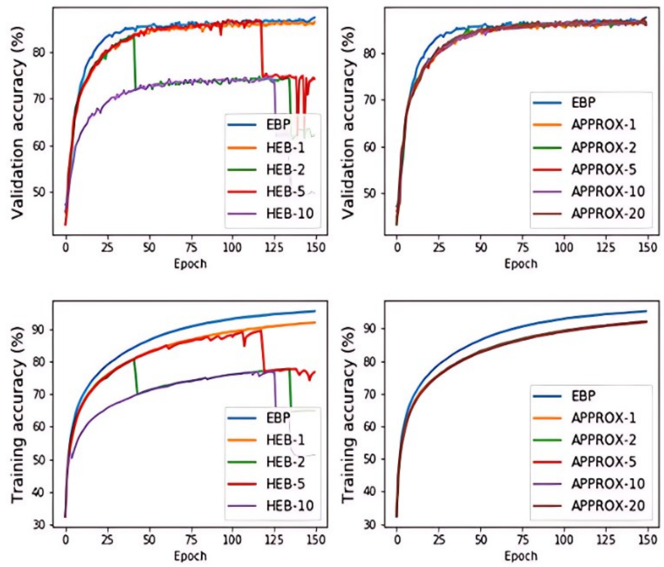


**Figure 7.** HEB and APPROX verification as well as performance on CIFAR4 for varied focus periods T, denoted as HEB-T and APPROX-T.

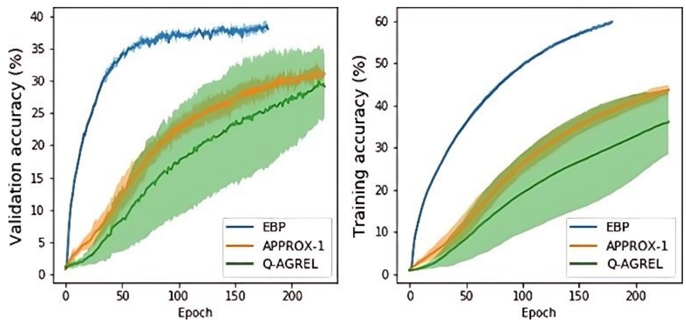


**Figure 8.** Validation and training accuracy of APPROX-1, EBP and Q-AGREL on CIFAR100.

**Table 1.** The MNIST patterns have been enhanced to target different focus periods during 'T' testing of WG, while also enhancing the reliability of HEB.

| | | T | α | Test accuracy | Epochs |
|---|---|---|---|---|---|
| MNIST | WG | 1 | 0.01 | 92.24 ± 0.41 | 79 ± 13 |
| | | 2 | 0.005 | 92.29 ± 0.38 | 83 ± 13 |
| | | 5 | 0.002 | 92.32 ± 0.42 | 80 ± 13 |
| | HEB | 1 | 0.01 | 98.48 ± 0.10 | 33 ± 8 |
| | | 2 | 0.005 | 98.45 ± 0.04 | 44 ± 21 |
| | | 5 | 0.002 | 80.85 ± 35.29 | 26 ± 14 |
| | EBP | - | 0.01 | 98.67 ± 0.04 | 29 ± 23 |

**Table 2.** Shifting focus spans 'T' on CIFAR10 and CIFAR4: HEB and APPROX quality. Bolded wording indicates the best outcomes.

| | | T | α | Test accuracy | Epochs |
|---|---|---|---|---|---|
| CIFAR10 | HEB | 1 | 0.005 | 71.05 ± 0.73 | 144 ± 7 |
| | | 2 | 0.0025 | 72.04 ± 0.19 | 140 ± 6 |
| | | 5 | 0.001 | 71.69 ± 0.82 | 138 ± 9 |
| | APPROX | 1 | 0.005 | 71.44 ± 0.51 | 142 ± 11 |
| | | 2 | 0.0025 | 71.45 ± 0.51 | 136 ± 14 |
| | | 5 | 0.001 | 72.34 ± 0.18 | 138 ± 5 |
| | EBP | - | 0.001 | 72.21 ± 0.64 | 134 ± 9 |
| CIFAR4 | HEB | 1 | 0.01 | 84.65 ± 0.76 | 124 ± 23 |
| | | 2 | 0.005 | 61.15 ± 29.52 | 96 ± 37 |
| | | 5 | 0.002 | 72.61 ± 23.81 | 115 ± 18 |
| | | 10 | 0.001 | 49.10 ± 29.51 | 99 ± 53 |
| | APPROX | 1 | 0.01 | 84.85 ± 0.30 | 131 ± 10 |
| | | 2 | 0.005 | 85.18 ± 0.50 | 122 ± 8 |
| | | 5 | 0.002 | 85.25 ± 0.90 | 118 ± 26 |
| | | 10 | 0.001 | 85.22 ± 0.69 | 130 ± 19 |
| | | 20 | 0.0005 | 84.98 ± 0.45 | 141 ± 10 |
| | EBP | - | 0.001 | 86.21 ± 0.51 | 134 ± 9 |

**Table 3.** Test accuracy of EBP, Q-AGREL and APPROX-1, on CIFAR100.

| | T | α | Test accuracy | Epochs |
|---|---|---|---|---|
| EBP | – | 0.001 | 41.16 ± 0.42 | 162 ± 11 |
| Q-AGREL | – | 0.01 | 32.38 ± 4.90 | 226 ± 2 |
| APPROX | 1 | 0.005 | 34.72 ± 0.96 | 217 ± 9 |

## 6. Result and Discussion

In this research, we looked at how well learning rules performed over five different network initializations, where the number of epochs taught varied according to the task complexity. For attention spans T = 1, 2, and 5, supervised error backpropagation (EBP) and the attention-modulated Hebbian rule (HEB) both beat the weight gain rule on MNIST. When it came to categorizing new data, HEB was competitive with EBP for attention spans of 1 and 2, but it rapidly declined for longer attention spans. Attention-guided learning (HEB) on CIFAR10 and its vehicle subset is compared to EBP for poor attention spans on basic tasks. For the fresh CIFAR10 data with attention spans T=1, 2, and 5, HEB marginally outperformed EBP. Rounding errors in the early layers made learning unreliable with T on the smaller set CIFAR4. In categorizing the fresh CIFAR10 data, AP-PROX minimizes these procedures and performs similarly or better than HEB.

No substantial attentional performance improvements are seen, demonstrating that neuronal output and attentional gain are linearly related to attention phase iterations. With backpropagation loss across all output nodes, APPROX-1 learns faster than Q-AGREL in categorizing training and fresh data. GRAD-CAM weights each engagement map that corresponds to predictive c significantly by summing the differences along its length and width, highlighting predicted zones. Every activated map's relevance is represented by the ac-weighted sum of the neural output A(x, y, k) in the final convolutional layer. To show how some regions have a beneficial effect on c, we employ ReLU. Figures 7 and 8 show GRAD-CAM-trained heat maps with prediction points highlighted. It is very unlikely that the enhanced focus is due to a decrease in cognitive percentage, given the concentration stage did not show any bias against important areas. The network correctly identified the 'bird' class in the first epoch and identified the proper label in the succeeding epochs, indicating a movement towards relevant locations throughout training.

## 7. Conclusion

Higher-performance computing (HPC) models play a crucial role in understanding and modifying systems, aiding in problem analysis, data collection, scenario creation, and input-driven behavior forecasting. Attention-guided learning and unusual information are growing in computational modeling, helping students reflect and understand systems. Transformer blocks employ parallelizability and less locally oriented attention than current or convolutional models. HPC software systems use attention-guided visual learning (AGVL), a computational paradigm that can compete with guided error backpropagation for classification problems. However, it is limited to competing over predicted categories using reinforcement signals. The study explores the use of transformer architectures to improve language modeling, focusing on attention-guided learning and attention-modulated Hebbian plasticity. Transformer architectures have an all-attention layer with embedded input vectors, non-contextual vectors with generic task-relevant information, and self-attentional and feedforward layers. The feature-similarity gain model suggests that attention multiplicatively makes neural responses stronger based on how similar their feature tuning is to the attended input. Supervised error backpropagation and the attention-modulated Hebbian rule both beat the weight gain rule on MNIST, but attention spans varied.

## Acknowledgments

We express our gratitude to International Research Institute from Bangladesh “RESIPO International” for providing the facilities and resources needed for this study. It was simpler to complete the study as intended with their help. We express our gratitude to our advisor, friends and colleagues for their encouragement, for fostering a lively academic environment, and for enabling us to work together on research projects.

## Data Availability Statement

Supplementary materials and data used in this research are accessible upon request for future research and development. We did not use any private, restricted, or licensed data in this research. For access, please contact the corresponding author via [*resipo.bd@gmail.com*].

## Additional Information

Authors did not receive funding for the research or publication. Both authors contributed equally, with technical and material support from "RESIPO International".

## Competing Interest Statement

The authors declare no known competing financial interests or personal relationships that could have influenced the work reported in this paper. Authors do not have any conflict of interests.

## References

[1] D. Baldauf and R. Desimone, “Neural mechanisms of object-based attention,” *Science*, vol. 344, no. 6182, pp. 424–427, 2014. https://doi.org/10.1126/science.1247003

[2] F. Baluch and L. Itti, “Mechanisms of top-down attention,” *Trends Neurosci.*, vol. 34, no. 4, pp. 210–224, 2011. https://doi.org/10.1016/j.tins.2011.02.003

[3] M. Carrasco, “Visual attention: the past 25 years,” *Vision Res.*, vol. 51, no. 13, pp. 1484–1525, 2011. https://doi.org/10.1016/j.visres.2011.04.012

[4] H. Qiu et al., “Boundary attention guided sparse feature learning for underwater object tracking in edge computing,” *ACM Trans. Multimed. Comput. Commun. Appl.*, 2024. https://doi.org/10.1145/3689824

[5] G. Dhara and R. K. Kumar, “Spatial attention guided cGAN for improved salient object detection,” *Front. Comput. Sci.*, vol. 6, 2024. https://doi.org/10.3389/fcomp.2024.1420965

[6] J. Bai et al., “UAV image object detection based on self-attention guidance and global feature fusion,” *Image Vis. Comput.*, vol. 151, no. 105262, p. 105262, 2024. https://doi.org/10.1016/j.imavis.2024.105262

[7] Z. A. Yazıcı et al., “GLIMS: Attention-guided lightweight multi-scale hybrid network for volumetric semantic segmentation,” *Image Vis. Comput.*, vol. 146, no. 105055, p. 105055, 2024. https://doi.org/10.1016/j.imavis.2024.105055

[8] W. Boulila, “An approach based on performer-attention-guided few-shot learning model for plant disease classification,” *Earth Sci. Inform.*, 2024. https://doi.org/10.1007/s12145-024-01339-x

[9] F. Yuan et al., “An effective CNN and Transformer complementary network for medical image segmentation,” *Pattern Recognit.*, vol. 136, no. 109228, p. 109228, 2023. https://doi.org/10.1016/j.patcog.2022.109228

[10] C. Yu and Y. Shin, “MCG-RTDETR: Multi-convolution and context-guided network with cascaded group attention for object detection in unmanned aerial vehicle imagery,” *Remote Sens. (Basel)*, vol. 16, no. 17, p. 3169, 2024. https://doi.org/10.3390/rs16173169

[11] O. Faruq et al., “Brain tumor MRI identification and classification using DWT, PCA and kernel support vector machine,” *Современные инновации, системы и технологии - Modern Innovations, Systems and Technologies*, vol. 4, no. 1, pp. 0133–0152, 2024. https://doi.org/10.47813/2782-2818-2024-4-1-0133-0152

[12] B. Ahmad et al., “DACB-Net: Dual attention guided compact bilinear convolution neural network for skin disease classification,” *arXiv* [cs.CV], 2024. https://doi.org/10.48550/ARXIV.2407.03439

[13] D. Oh et al., “GCN-assisted attention-guided UNet for automated retinal OCT segmentation,” *Expert Syst. Appl.*, vol. 249, no. 123620, p. 123620, 2024. https://doi.org/10.1016/j.eswa.2024.123620

[14] R. T. Hafiza et al., “Analysis of thyroid nodule ultrasound images by image feature extraction technique,” *Современные инновации, системы и технологии - Modern Innovations, Systems and Technologies*, vol. 4, no. 3, pp. 0301–0325, 2024. https://doi.org/10.47813/2782-2818-2024-4-3-0301-0325

[15] M. Ullah, I. A. Taj, and R. H. Raza, “Degradation model and attention guided distillation approach for low resolution face recognition,” *Expert Syst. Appl.*, vol. 243, no. 122882, p. 122882, 2024. https://doi.org/10.1016/j.eswa.2023.122882

[16] M. Abdar et al., “A review of uncertainty quantification in deep learning: Techniques, applications and challenges,” *Inf. Fusion*, vol. 76, pp. 243–297, 2021. https://doi.org/10.1016/j.inffus.2021.05.008

[17] I. M. Jahidul and O. Faruq, “Further exploration of deep aggregation for shadow detection,” *Современные инновации, системы и технологии - Modern Innovations, Systems and Technologies*, vol. 2, no. 3, pp. 0312–0330, 2022. https://doi.org/10.47813/2782-2818-2022-2-3-0312-0330

[18] Z. Zhan et al., “AMFNet: Attention-guided multi-scale fusion network for bi-temporal change detection in remote sensing images,” *Remote Sens. (Basel)*, vol. 16, no. 10, p. 1765, 2024. https://doi.org/10.3390/rs16101765

[19] Y. Wen, W. Ke, and H. Sheng, “Camouflaged object detection based on deep learning with attention-guided edge detection and multi-scale context fusion,” *Appl. Sci. (Basel)*, vol. 14, no. 6, p. 2494, 2024. https://doi.org/10.3390/app14062494

[20] O. Faruq, I. M. Jahi, M. S. Ahmed, and M. S. Hossain, “Brain tumor MRI identification and classification using DWT, PCA, and KSVM,” *Research Square*, 2023. https://doi.org/10.21203/rs.3.rs-2562932/v1

[21] O. Faruq, M. J. Islam, M. S. Ahmed, and M. S. Hossain, “Brain tumor MRI identification and classification using DWT, PCA, and KSVM,” *TechRxiv*, 2023. https://doi.org/10.36227/techrxiv.21771329.v2

[22] L. Zheng, Y. Cui, S. Jin, and Y. Chen, “High-performance computing-based open-source power transmission and distribution grid co-simulation,” *IEEE Trans. Power Syst.*, vol. 39, no. 5, pp. 6144–6153, 2024. https://doi.org/10.1109/tpwrs.2023.3349200

[23] S. Ning et al., “Photonic-electronic integrated circuits for high-performance computing and AI accelerators,” *J. Lightwave Technol.*, pp. 1–26, 2024. https://doi.org/10.1109/jlt.2024.3427716

[24] S. Wang, H. Zheng, X. Wen, and S. Fu, “Distributed high-performance computing methods for accelerating deep learning training,” *J. Knowl. Learn. Sci. Technol.*, vol. 3, no. 3, pp. 108–126, 2024. https://doi.org/10.60087/jklst.v3.n3.p108-126

[25] A. Pousa, V. M. Sanz, M. Naiouf, and A. De Giusti, “Trends in high performance computing and quantum computing,” *J. Comput. Sci. Technol.*, vol. 24, no. 2, p. e11, 2024. https://doi.org/10.24215/16666038.24.e11

[26] H. Zhang et al., “Multi-cluster high performance computing method based on multimodal tensor in enterprise resource planning system,” *Phys. Commun.*, vol.

62, no. 102231, p. 102231, 2024. https://doi.org/10.1016/j.phycom.2023.102231

[27] H. Xu, S. Song, and Z. Mao, “Characterizing the performance of emerging deep learning, graph, and high-performance computing workloads under interference,” in *2024 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2024, pp. 468–477. https://doi.org/10.1109/ipdpsw63119.2024.00098

[28] J. Yang et al., “Parallel algorithm design and optimization of geodynamic numerical simulation application on the Tianhe new-generation high-performance computer,” *J. Supercomput.*, vol. 80, no. 1, pp. 331–362, 2024. https://doi.org/10.1007/s11227-023-05469-9