Training-Free Approach of Convolutional Neural Networks with Astrocyte-Inspired Architectures

Ana Ribas-Rodriguez
Research Center on ICT
Computer Science and Information Technology Dept.
University of A Coruña
A Coruña, Spain
ana.ribas@udc.es

Vanessa Aguiar-Pulido
Dept. of Computer Science
University of Miami
Coral Gables, FL, USA
vanessa@cs.miami.edu

Francisco Cedron
Research Center on ICT
Computer Science and Information Technology Dept.
University of A Coruña
A Coruña, Spain
francisco.cedron@udc.es

Abstract

This work introduces the Artificial Neuron-Astrocyte Network (ANAN), a novel approach that incorporates artificial astrocytes into pre-trained Convolutional Neural Networks (CNN) to enhance performance without requiring additional training. By dynamically modulating synaptic weights based on neuronal activity, astrocytes allow the network to adapt to input data efficiently. The proposed approach only requires optimizing four parameters, instead of the millions that are typically required in CNN fine-tuning. This offers a resource-saving alternative to traditional fine-tuning methods. Experimental results demonstrate statistically significant improvements in performance employing four different datasets, including one balanced and one imbalanced biomedical dataset, as well as two balanced ones encompassing natural images. Results in different application domains highlight the potential of astrocytes to optimize network performance without the need of going through traditional training cycles.

1 Introduction

Artificial neural networks (ANNs) emerged from trying to gain a deeper understanding of how the brain works. The first computerized model of an artificial neuron was introduced by Warren McCulloch and Walter Pitts in 1943 and, since then, ANNs have been constantly evolving and improving their performance [1]. Neurons are an important part of the nervous system, but they are not the only relevant cells. Santiago Ramón y Cajal, who can be considered as the father of modern neuroscience, had already observed the presence of glial cells in his research. He noticed that these cells are abundant and ubiquitous in the brain, and he hypothesized they may play an important role [2]. Indeed, as science advanced, it was proven that astrocytes, a type of glial cell, actively participate in information processing alongside neurons, in what is known today as the tripartite synapse [3]. Figure 1 provides an overview of the tripartite synapse, the main inspiration for this work. In a tripartite synapse, an astrocyte modulates the transmission of information between a pre-synaptic neuron and a post-synaptic one. These discoveries highlight the biological importance of astrocytes, which were once thought to have a passive, supporting role of neurons (e.g., routine housekeeping duties) [4].

It is worth noting that astrocyte activity happens on a different timescale than that of neurons, ranging from seconds to hours [5]. Thus, neurons respond swiftly to stimuli, whereas astrocytes operate on slower timescales [6] and they modulate both excitatory and inhibitory transmission between neurons [7]. Another well-known property of astrocytes and glial cells is that the ratio between glia and neurons increments as brain size does in the phylogenetic scale, suggesting a correlation between

Latinx in AI @ NeurIPS-24 (LXAI 2024).

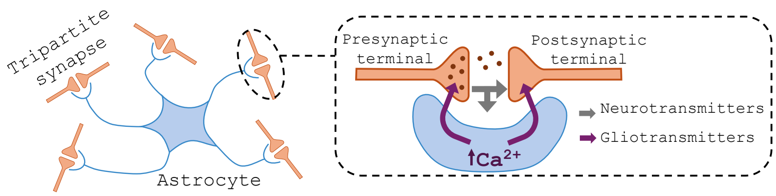

Figure 1: The tripartite synapse involves an astrocyte, a pre-synpatic neuron and a post-synaptic one. An astrocyte with its prolongations wrapping multiple synapses is shown on the left; on the right, a tripartite synapse schematic shows pre- and post-synaptic terminals. Astrocytes capture neurotransmitters and release gliotransmitters.

these cells and cognition [8]. Exploring this idea, and given that astrocytes have been described as affecting different behavioral aspects (like cognition) in mice, Windrem et al. [9], and Han et al. [10] introduced human astrocytes into mice brains. The experimental results show that this leads to an improvement of their cognitive abilities.

In this work, we propose a novel training-free approach in which we add artificial astrocytes to a pre-trained Convolutional Neural Network (CNN). In the proposed Artificial Neuron-Astrocyte Network (ANAN), artificial astrocytes modulate neuronal synapses by adjusting the weights based on a neuronal activation threshold, reinforcing the positive (excitatory) and negative (inhibitory) response. The proposed approach is tested on four different datasets across two different domains of application (natural images and biomedical images). Balanced and imbalanced datasets were used to assess the behavior of the proposed approach under different conditions. The ANAN is compared to a base network that has been previously trained on each dataset. Results suggest that the addition of artificial astrocytes post-training enhances the performance of CNNs.

This work attempts to replicate the biological phenomenon seen in the work by Windrem et al. [9], and Han et al. [10]. Thus, the already trained neural network would correspond to the mouse brain, and the new processing elements (i.e., artificial astrocytes) would simulate the behavior of human astrocytes. Each input (i.e., an image) is passed multiple times to simulate the difference in timescale between neurons and astrocytes. By comparing an already existing network with and without astrocytes added after training, we aim to study whether the computational model is capable of reproducing the results observed in the biological model (i.e., an improvement of cognitive abilities). The novelty of this work lies in the introduction of astrocytes to deep models used for computer vision tasks without requiring retraining, thus using less resources. Introducing these elements post-training allows the CNN to generalize and adapt to a different domain of application.

2 Related work

In recent years, the field of neuroscience has become less neuron-centric, shifting its focus toward the study of glial cells. Experiments have shown that the release of neurotransmitters is modulated by glial cells, specifically astrocytes, in what is called the tripartite synapse. Since the discovery of the tripartite synapse [3], methods have emerged attempting to translate this concept into *in silico* models by incorporating astrocytes into classical ANNs. The first published works differ in the ANN architecture employed, as well as in the quantity and location of the astrocytes. While Ikuta et al. [11] incorporated one astrocytic layer in a multilayer perceptron (MLP), Mesejo et al. [12] included one astrocyte per neuron in all of the neurons of an MLP. Landolsi and Marzouki [13] employed a SONG-NET architecture and Porto-Pazos et al. [14] added the astrocytes to an MLP using evolutionary algorithms. In their work, Porto-Pazos et al. [14] established that the improvement in performance of the network is specifically due to astrocytes. Moreover, these authors show that this improvement increases as the network complexity increases, and that the impact of the astrocytes depends on the problem tested. In all these works, the ANNs were trained with the astrocytes already incorporated [15].

While early research of astrocyte integration into neural networks was carried out using classical models, recent methods have modeled more realistic astrocyte-neuron interactions in the brain. One such example is the incorporation of astrocytes into spiking neural networks (SNNs). These networks are believed to be more biologically plausible and enhance processing efficiency compared to traditional networks [16–19]. Moreover, several studies show that astrocytes that have been added to deep learning SNN-like architectures can significantly improve performance in a variety of tasks. For instance, He et al. [18] state that including astrocytes to update neuron states in SNNs increased accuracy when using image, voice, and neural morphology datasets. Stasenko and Kazantsev [19] found that astrocytes help suppress noise and improve image representation in SNNs, acting as a buffer for processing information patterns. Similarly, Nazari et al. [17] demonstrated that combining SNNs with astrocytes improved pattern recognition and facilitated information transmission. The Blended Glial Cell's Spiking Neural Network proposed by Tao et al. [16] incorporates astrocytes as spatiotemporal information processing units, enhancing neuron and synapse plasticity. Their network, tested in applications like Sudoku solving, showed superior accuracy and computational efficiency compared to other solvers, emphasizing the benefits of adding astrocytes.

In addition to the above, other authors have created more complex architectures featuring astrocytes. Han et al. [20] introduced AstroNet, which uses astrocytes to optimize neuronal connections and improve accuracy. MA-Net, a more advanced version of the previous approach, modulates connections during training, achieving state-of-the-art performance with fewer parameters [21]. Zimin et al. [22] proposed a hybrid ANN that incorporates astrocyte-driven short-term memory, significantly improving performance in visual tasks compared to traditional models. The authors of [23] further showed that neuron-astrocyte networks enhance associative memory capacity and efficiency, supporting the crucial role of astrocytes in memory storage in an energy-based model. Finally, Kozachkov et al. [24] hypothesized that neuron-astrocyte networks implement the core computation performed in the transformer block in artificial intelligence (AI) (i.e., the normalization operation in the self-attention) and provide a potential biological explanation of how transformers relate to the brain.

The majority of these studies design their own (complex) architecture to include astrocytes in ANNs. All the models described in this section have a training phase in which the astrocytes are trained while training the model. Astrocytes are shown to be problem dependent so, like ANNs, their training phase may be costly in terms of time and memory. Training-free approaches are particularly attractive in environments with limited resources, as the number of parameters to be optimized is much smaller. Studies adding astrocytes to existing pre-trained networks are lacking. Hence, it would be interesting to explore how astrocytes work in an already pre-trained model, mimicking the biological experiments by Windrem et al. [9] and Han et al. [10], in which the addition of human astrocytes to mice increases their cognitive abilities. To the best knowledge of the authors, no published work explores this possibility. This work, thus, studies how adding astrocytes to a pre-trained network affects its performance.

3 Materials and methods

3.1 Datasets

Four datasets were used to evaluate the performance of the proposed approach in binary classification. The **Cats&Dogs** dataset contains images of cats and dogs. The dataset is not the classical, larger one, but a reduced version of it that was obtained from Kaggle [25]. This dataset was originally assembled from images retrieved from Google. The **Fire** dataset includes natural images with and without wildfire. It was retrieved from Kaggle [26] and the photos were originally downloaded from Google. The **RFMID** dataset comprises images related to various ophthalmic diseases [27]. In this work, only the images corresponding to diabetic retinopathy (not healthy) and healthy individuals were used. Hence, the problem was simplified from a multi-class classification to a binary classification of healthy versus not healthy. The total number of images employed is shown in Table 1. Finally, the **RMNIST** dataset was extracted from MedMNIST [28] but originally the data is from a study by Liu et al. [29]. In this work, we only distinguish between healthy/not healthy, however originally the dataset contained information pertaining to different diabetic retinopathy disease stages.

A summary of the characteristics of the four datasets is presented in Table 1. All the datasets employed in this study are balanced except for the Fire dataset. Since we are using the already partitioned test set made by the original authors, the ratio of test images varies slightly across datasets. The Cats&Dogs and Fire test sets employed nearly 20% of the total number of images for testing, while the RFMID and the RMNIST test sets encompassed 24.7% and 33.33% of the total number of images respectively.

Table 1: Binary Datasets Distribution

	Cats&Dogs	Fire	RFMID	RMNIST
Class 0 (%)	50.07	75.58	51.42	44.62
Class 1 (%)	49.93	24.42	48.58	55.38
# images	697	999	1031	1600

3.2 Base network

First, a traditional CNN was designed using TensorFlow [30], with the goal of having a baseline with which to compare the proposed approach. The network architecture is shown in Figure 2. Slight modifications were introduced to the base network to achieve better performance:

- The network used for the RFMID and RMNIST datasets includes an additional layer and uses 5x5 filters instead of 3x3 due the nature of the images included in these datasets.
- The input image size varies across datasets: 224x224 pixels for the RFMID, RMNIST and Fire datasets, and 128x128 pixels for the Cats&Dogs dataset.

Despite these variations, the overall architectures of the four models are quite similar. All CNN models were trained using 10-fold cross-validation for a maximum of 100 epochs and a batch size of 64, with ReLU as the activation function and Adam as the optimizer. The stopping criterion was based on the loss; training was halted if the loss did not change by 0.01 units over 30 consecutive epochs. After training, artificial astrocytes were added to the first convolutional layer of each network.

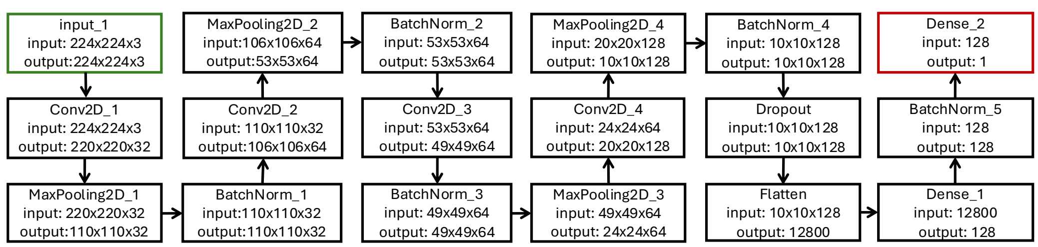

Figure 2: Convolutional Neural Network architecture used as a basis. The input layer is shown in green, while the output layer is depicted in red.

3.3 Artificial Neuron-Astrocyte Network (ANAN)

3.3.1 General overview

Several studies mentioned in Section 2 show that astrocytes can filter noisy cues [31, 32, 19]. Additionally, the works by Windrem et al. [9] and Han et al. [10] suggest that adding human astrocytes to mice enhances their cognitive capabilities. We sought to translate these biological phenomena into a computational model by incorporating artificial astrocytes post-training, something that, to the best knowledge of the authors, has not been done before. Thus, in this work, we added artificial astrocytes to the first convolutional layer of the already trained CNN described in Section

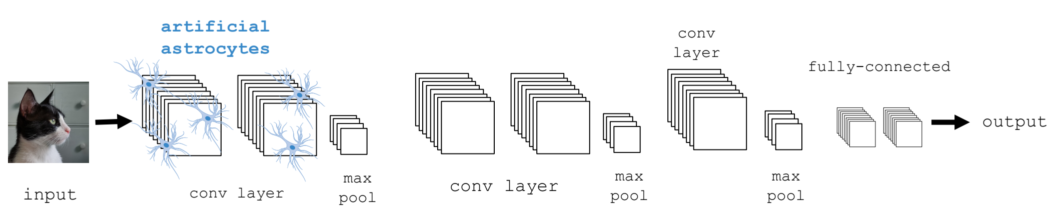

Figure 3: Overview of the Artificial Neuron-Astrocyte Network (ANAN). Astrocytes are added to the first convolutional layer of the base CNN and each of them modulates the pre-synaptic weights (filters) associated with each kernel in this layer.

3.2. An overview of the proposed approach - the Artificial Neuron-Astrocyte Network (ANAN) - is depicted in Figure 3.

Subsequently, the steps followed to process an input (image) are described. Astrocytes modify the base network, but this modification only lasts while the current image is presented. Thus, first, a copy of the initial state of the network is saved (i.e., the CNN model with its weights). Next, an input image is passed to the ANAN. Each image is presented during a number of iterations to the network. During this time, if activated, the astrocytes modify the network's weights. Finally, after obtaining the prediction for the current input, the network's weights are restored to their original values. Therefore, the effect of the astrocytes is temporary. The process described is depicted in Figure 4.

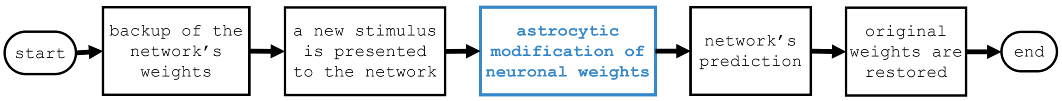

Figure 4: Workflow showing the behavior of the ANAN for a single input which includes (from left to right): (1) backing up the network's weights, (2) presenting an image to the network, (3) modifying the neuronal weights according to the astrocytic algorithm (described in more detail in Section 3.3.2), (4) obtaining the network's prediction for the input image, and (5) restoring the network to its original state.

3.3.2 Astrocytic modulation

As previously mentioned, astrocytes act as modulators of neural transmission at the synaptic cleft, similarly to their biological counterparts in tripartite synapses. They modify the pre-synaptic weights of neurons (kernels) in the first convolutional layer. This modification can be either excitatory (information transmission is potentiated) or inhibitory (information transmission is attenuated), similarly to how the information is potentiated or inhibited along the neural pathways in the brain. Furthermore, biological astrocytes operate on a slower timescale than neurons. To reflect this in our computational model, a single stimulus (image) is passed to the network during a specific number of iterations. During this time, the astrocytes monitor the output of the neurons in the first convolutional layer. Depending on the number of times the post-synaptic neuron (kernel) is activated, the pre-synaptic weights (filters) are updated. Below the exact process followed to update the weights is described in detail.

The ANAN features the following astrocytic parameters: $\boldsymbol{k}$, $\boldsymbol{\mu}$, $\boldsymbol{a}$, and $\boldsymbol{b}$.

- k is the number of iterations during which a single image is presented (simulating the slower astrocytic communication).
- μ acts as the trigger that activates the astrocytic effect. It can have an excitatory (μ) or inhibitory effect ($-\mu$).
- a and b represent the values by which the weights are incremented or decreased, respectively.

These astrocytic parameters behave as described subsequently.

The artificial astrocytes monitor the neuron's activity (i.e., the output or $y_j(t)$), just as the biological astrocytes in the tripartite synapse do, during k iterations. To do so, function $u : \mathbb{R} \to \mathbb{Z}$ is applied to the neuron's output. This function simulates the release of neurotransmitters by the biological neuron and indicates whether the neuron has been activated or not, and it is defined as follows:

$$u(x) = \begin{cases} -1 & x \leq 0 \\ 1 & x > 0 \end{cases} \tag{1}$$

Afterward, the astrocyte modifies the pre-synaptic weights of the neuron when the trigger (activation level) reaches a threshold μ or $-\mu$, depending on whether it is positive (excitatory) or negative (inhibitory). Function $r_j : \mathbb{N} \setminus \{0\} \to [-\mu, \mu]$ outputs the number of times a neuron has been activated and is defined as:

$$r_j(t) = \begin{cases} u(y_j(t)) + r_j(t-1) & t > 0, \quad r_j(t-1) \in (-\mu, \mu) \\ r_j(t-1) & t > 0, \quad r_j(t-1) \in \{-\mu, \mu\} \\ u(y_j(t)) & \text{otherwise} \end{cases} \tag{2}$$

Once a neuron j reaches the activation level μ (or $-\mu$), the astrocyte modifies its associated pre-synaptic weights w_i. Astrocytic adjustment of these pre-synaptic weights is defined as $w_i(t + \Delta t) = w_i(t) + \Delta w_i(t)$, where $\Delta w_i(t) = |w_i(t)|z(t)$ and $z : \mathbb{N} \setminus \{0\} \to \mathbb{R}$ is a function defined as follows:

$$z(t) = \begin{cases} a & r_j(t) = \mu \\ -b & r_j(t) = -\mu \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

Hence, if the neuron is activated μ times, the pre-synaptic weights will be increased by a percentage a, but they will be decreased by a percentage b if the neuron is activated $-\mu$ times. When a neuron reaches the activation levels $\{-\mu, \mu\}$, the trigger value is maintained and does not reset as a consequence of astrocytic excitation. Consequently, the astrocytic effect lasts all k iterations, reinforcing the connection over time as described in ref. [12].

To sum up, the pre-synaptic weights are modified when the astrocyte is activated by a neuron firing. In the next iteration, if the neuron keeps firing, it further activates the astrocyte and thus the weights are modified again. Once the stimulus is not present anymore and before the next stimulus (image) is presented, the weights are reset back to their original state.

3.3.3 Hyperparameter optimization

Four parameters (k, μ, b, a) determine the behavior of the astrocytes in the network, and the optimal values were found through grid search, as they vary depending on the problem to solve. The range of parameters employed for the grid search is chosen based on previous work by Porto-Pazos et al. [14]:

- For k: Values between 4 and 8 were tested.
- For μ: Values between 2 and 4 were tested.
- For a: Values between 1.05 and 1.50 were tested, with an increase of 0.15.
- For b: Values between 0.05 and 0.50 were tested, with an increase of 0.15.

4 Results

A grid search was performed to obtain the optimal configuration of the astrocytic parameters (k, μ, b, a) of the ANAN, as described in the previous section. Table 2 displays the best parameters obtained for each dataset.

Table 2: Best configuration of astrocytic parameters for each dataset.

Parameter	Cats&Dogs	Fire	RFMID	RMNIST
k	4	4	8	5
μ	2	2	3	4
a	1.05	1.50	1.05	1.05
b	0.30	0.50	0.30	0.05

The results obtained for each dataset used are presented in Table 3. This table allows comparing the performance of the base CNN with the proposed approach (i.e., ANAN), using metrics such as accuracy, F1-Score and and Binary-Cross-Entropy loss. Note that while the loss metric is generally used in contexts in which a network is trained, in this case it is only used as a measure of each model's performance on the test set. The values shown in the table correspond to the application of the base CNN and the ANAN to the test set for each dataset. Significance is shown where appropriate in addition to performance metrics, highlighting the best value in bold. To calculate statistical significance (i.e., p-values) between the performance of the base CNN and the ANAN, the non-parametric Wilcoxon statistical test was used [33]. By comparing a network with and without astrocytes, we are able to assess their impact on performance. It is worth noting that the addition of astrocytes does not entail retraining the deep network.

The results show that for the **Cats&Dogs** dataset, the ANAN model achieved a significantly better loss compared to the base CNN, with p<0.01, and although the F1-score was higher for the ANAN, it was not statistically significant. For the **Fire** dataset, the accuracy and the F1-score were both significantly better for the ANAN with $p<0.05$, while the loss was significantly worse with p<0.01. For the **RFMID** dataset, the F1-score was significantly better for the ANAN with $p<0.01$, but the loss was significantly worse for the ANAN with p<0.05. Finally, for the **RMNIST**, the ANAN model was significantly better in terms of F1-score, with $p<0.05$, the accuracy was higher for the ANAN but not significantly, and the loss was significantly worse for the ANAN. It is worth highlighting that the

Table 3: Results obtained using a Convolutional Neural Network (CNN) and the proposed Artificial Neuron-Astrocyte Network (ANAN) for four different datasets. Statistical significance levels are shown with asterisks: ** $p < 0.01$, * $p < 0.05$ next to the best performing algorithm.

	Cats&Dogs		Fire		RFMID		RMNIST	
Metric	CNN	ANAN	CNN	ANAN	CNN	ANAN	CNN	ANAN
Loss	16.127	**1.354** **	**0.066** **	0.168	**0.529** *	0.610	**1.243** **	1.674
Accuracy	**0.662**	0.651	0.981	**0.987** *	**0.865**	0.859	0.754	**0.765**
F1-Score	0.613	**0.668**	0.959	**0.973** *	0.737	**0.843** **	0.765	**0.785** *

base CNN was trained using the loss as a stopping criterion; hence, in general, we would expect the loss values to be smaller for the base CNN as this model has been optimized for this metric. Unlike the CNN, the ANAN does not go through a traditional training process.

The F1-Score provides a measure of a model's performance in terms of precision and recall (i.e., the F1-Score is the harmonic mean of precision and recall). This is one of the preferred metrics to evaluate the difference in network performance when using imbalanced datasets. Given the imbalance observed in the Fire dataset, we decided to analyze the results obtained in terms of F1-Score further. Figure 5 illustrates the differences in terms of F1-score and accuracy across all four datasets. While accuracy values are quite similar between both approaches, we can see that the ANAN consistently outperforms the CNN in terms of F1-score, with statistically significant improvements for three datasets (RFMID, Fire and RMNIST), but not for the Cats&Dogs dataset. The later could be explained by the wider range of values obtained for the base CNN.

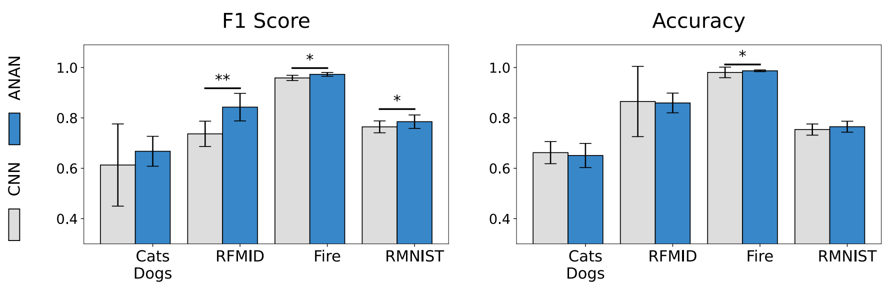

Figure 5: Comparison between the base Convolutional Neural Network and the same network adding astrocytes (i.e., the proposed approach, ANAN). F1-score and accuracy metrics are shown for the different datasets used. Significance level is shown with asterisks: ** $p < 0.01$, * $p < 0.05$.

5 Conclusion and Future work

In this work, we propose a novel approach, the Artificial Neuron-Astrocyte Network (ANAN), that introduces artificial astrocytes in the first convolutional layer of a pre-trained CNN. The proposed approach was tested on four datasets, comparing the performance of a baseline CNN with the ANAN in binary classification. Results highlight the potential of astrocytes to optimize network performance. All images used in this work are publicly available and those of biomedical origin were de-identified by the original authors, thus limiting any concerns that may arise regarding privacy and security.

The primary innovation of the proposed approach lies in the fact that it leverages a pre-trained CNN and incorporates artificial astrocytes to enhance its performance without requiring additional network training. These artificial astrocytes dynamically adjust the pre-synaptic weights of the first convolutional layer based on the neuronal (kernel) activity observed over a specified period of time (i.e., a number of iterations), doing so in a controlled, time-sensitive manner that modifies the network's behavior as it processes each input. This unique mechanism enables the network to adapt to input data effectively, enhancing its performance without the extensive retraining that traditional fine-tuning methods require. Results suggest that introducing artificial astrocytes improves network performance in terms of F1-score across all the datasets used. Although the tests carried out here involve binary classification only to simplify the initial evaluation, the methodology could be seamlessly applied in the future to multi-class classification. The core principles of astrocyte-driven weight modulation remain consistent across various classification tasks, ensuring the framework's versatility.

Astrocytic weight modulation differs fundamentally from conventional fine-tuning in terms of computational cost. Typically, fine-tuning involves optimizing a wide range of parameters, which can be computationally intensive, and the number of variables can easily reach millions or even billions. In contrast, the ANAN utilizes a fixed set of four parameters for astrocytic modulation of the weights, achieving good results with fewer variables to manage, thereby reducing computational costs compared to traditional fine-tuning methods. This efficiency is a key benefit of the proposed approach, aligning with the growing emphasis on sustainable AI practices [34] and which could offer a streamlined and resource-efficient alternative. This may be particularly valuable for deploying models in production settings, where fine-tuning is often costly. By leveraging a pre-trained CNN as its foundation, the ANAN bypasses the additional training cycles typically required by traditional methods, potentially limiting the amount of computational resources required and aligning with eco-friendly AI practices. Approaches like this one may be the key to reducing the environmental impact of deep learning models in future research, potentially offering a more sustainable path for AI development.

Astrocytes, active in a context-dependent manner, can reconfigure neural pathways on demand serving as the natural biological substrate for context representation in ANNs. They also represent a slower information flow in the brain, integrating important information from various sources across their extensive, non-overlapping networks [35]. As discussed in ref. [36], including this behavior in ANN-like models might enable more accurate and robust modeling for complex tasks. This could, in turn, open new avenues for these astrocyte-inspired networks in deep learning, representing a significant advancement in promoting both efficiency and sustainability. Conversely, modeling astrocytic behavior using computational models may yield discoveries in neuroscience, particularly regarding the role of biological neuron-astrocyte networks in memory storage and consciousness, and could also provide insights into the framework for biological astrocytic function proposed by Murphy-Royal et al. [36].

Future work will include testing the proposed approach on multiple multi-class datasets across different application domains, as well as employing more complex networks (e.g., VGG, ResNet, ConvNeXT) as a basis. As noted by Porto-Pazos et al. [14], the benefits of astrocytes become increasingly apparent as the complexity of neural architectures grows, mirroring the biological trend where the proportion of glial cells per neuron increases with nervous system complexity [8]. This suggests that the ANAN's impact may be even more pronounced as deep learning models grow in complexity. Moreover, refining the astrocytic parameter space could further reduce the time required for the ANAN to converge, resulting in faster model convergence.

Another avenue of research will involve optimizing the position of astrocytes in the network as an astrocytic parameter to achieve better performance, an area that, to the best knowledge of the authors, has not yet been explored. Modeling astrocytes as an interconnected network as some researchers have already done [18, 21, 17, 16] may possibly be the next step towards more biologically accurate models. Furthermore, implementing astrocytes as a network for short-term memory buffering, as suggested by Tsybina et al. [31] and Zimin et al. [37], could help improve performance. This could eventually support or yield some interesting theories about how memories are stored in the brain, as Robertson [38] proposes, suggesting that explicit memories could be stored and encoded in astrocyte extensive tessellating domains within the neocortex. In addition to classification, the proposed approach could be employed for other tasks such as regression, segmentation or reinforcement learning. Finally, the idea behind the ANAN could also be extended to natural language processing, real-time video analysis, or introduced in large language models, for example.

6 Acknowledgement

We are particularly grateful to CESGA (Galician Supercomputing Centre) for providing access to their infrastructure. The work of Vanessa Aguiar-Pulido is supported by a Provost Research Award from the University of Miami.

References

[1] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133, December 1943. ISSN 1522-9602. doi: 10.1007/bf02478259. URL `http://dx.doi.org/10.1007/BF02478259`.

[2] Virginia García-Marín, Pablo García-López, and Miguel Freire. Cajal's contributions to glia research. *Trends in neurosciences*, 30(9):479–487, 2007.

[3] Alfonso Araque, Vladimir Parpura, Rita P Sanzgiri, and Philip G Haydon. Tripartite synapses: glia, the unacknowledged partner. *Trends in neurosciences*, 22(5):208–215, 1999.

[4] Alexei Verkhratsky, Margaret S Ho, Robert Zorec, and Vladimir Parpura. The concept of neuroglia. *Adv. Exp. Med. Biol.*, 1175:1–13, 2019.

[5] Nina Vardjan, Vladimir Parpura, and Robert Zorec. Loose excitation-secretion coupling in astrocytes. *Glia*, 64(5):655–667, May 2016.

[6] Gertrudis Perea, Marta Navarrete, and Alfonso Araque. Tripartite synapses: astrocytes process and control synaptic information. *Trends in Neurosciences*, 32(8):421–431, August 2009. ISSN 0166-2236. doi: 10.1016/j.tins.2009.05.001. URL `http://dx.doi.org/10.1016/j.tins.2009.05.001`.

[7] Tommaso Fellin. Communication between neurons and astrocytes: relevance to the modulation of synaptic and network activity. *J. Neurochem.*, 108(3):533–544, February 2009.

[8] Nicola J Allen and Ben A Barres. Neuroscience: Glia - more than just brain glue. *Nature*, 457 (7230):675–677, February 2009.

[9] Martha S. Windrem, Steven J. Schanz, Carolyn Morrow, Jared Munir, Devin Chandler-Militello, Su Wang, and Steven A. Goldman. A competitive advantage by neonatally engrafted human glial progenitors yields mice whose brains are chimeric for human glia. *The Journal of Neuroscience*, 34(48):16153–16161, November 2014. ISSN 1529-2401. doi: 10.1523/jneurosci.1510-14.2014. URL `http://dx.doi.org/10.1523/JNEUROSCI.1510-14.2014`.

[10] Xiaoning Han, Michael Chen, Fushun Wang, Martha Windrem, Su Wang, Steven Shanz, Qiwu Xu, Nancy Ann Oberheim, Lane Bekar, Sarah Betstadt, Alcino J Silva, Takahiro Takano, Steven A Goldman, and Maiken Nedergaard. Forebrain engraftment by human glial progenitor cells enhances synaptic plasticity and learning in adult mice. *Cell Stem Cell*, 12(3):342–353, March 2013.

[11] Chihiro Ikuta, Yoko Uwate, and Yoshifumi Nishio. Chaos glial network connected to multi-layer perceptron for solving two-spiral problem. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pages 1360–1363, 2010. doi: 10.1109/ISCAS.2010.5537060.

[12] Pablo Mesejo, Oscar Ibánez, Enrique Fernández-Blanco, Francisco Cedrón, Alejandro Pazos, and Ana B Porto-Pazos. Artificial neuron–glia networks learning approach based on cooperative coevolution. *International journal of neural systems*, 25(04):1550012, 2015.

[13] Hajer Landolsi and Kirmene Marzouki. Self organizing neuro-glial network, song-net. In *2014 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*, pages 85–91, 2014. doi: 10.1109/CCMB.2014.7020698.

[14] Ana B. Porto-Pazos, Noha Veiguela, Pablo Mesejo, Marta Navarrete, Alberto Alvarellos, Oscar Ibáñez, Alejandro Pazos, and Alfonso Araque. Artificial astrocytes improve neural network performance. *PLOS ONE*, 6(4):1–8, 04 2011. doi: 10.1371/journal.pone.0019109. URL `https://doi.org/10.1371/journal.pone.0019109`.

[15] Sara Alvarez-Gonzalez, Francisco Cedron, Alejandro Pazos, and Ana B Porto-Pazos. Artificial glial cells in artificial neuronal networks: a systematic review. *Artificial Intelligence Review*, 56 (Suppl 2):2651–2666, 2023.

[16] Liying Tao, Pan Li, Meihua Meng, Zonglin Yang, Xiaozhuang Liu, Jinhua Hu, Ji Dong, Shushan Qiao, Tianchun Ye, and Delong Shang. Blended glial cell's spiking neural network. *IEEE Access*, 11:43566–43582, 2023. doi: 10.1109/ACCESS.2023.3267856.

[17] Soheila Nazari, Masoud Amiri, Karim Faez, and Marc M. Van Hulle. Information transmitted from bioinspired neuron–astrocyte network improves cortical spiking network's pattern recognition performance. *IEEE Transactions on Neural Networks and Learning Systems*, 31(2): 464–474, 2020. doi: 10.1109/TNNLS.2019.2905003.

[18] Junwen He, Jing Fan, Yuhan Zheng, Yunchuan Liu, Hong-Yu Zhang, and Hui Peng. Gcsnn: Glia-coupled spiking neural network. In *2023 China Automation Congress (CAC)*, pages 9161–9166, 2023. doi: 10.1109/CAC59555.2023.10452072.

[19] Sergey V Stasenko and Victor B Kazantsev. Dynamic image representation in a spiking neural network supplied by astrocytes. *Mathematics*, 11(3):561, 2023.

[20] Mengqiao Han, Liyuan Pan, and Xiabi Liu. Astronet: When astrocyte meets artificial neural network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20258–20268, 2023.

[21] Mengqiao Han, Liyuan Pan, and Xiabi Liu. Ma-net: Rethinking neural unit in the light of astrocytes. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(3):2040–2048, Mar. 2024. doi: 10.1609/aaai.v38i3.27975. URL `https://ojs.aaai.org/index.php/AAAI/article/view/27975`.

[22] Ilya A. Zimin, Victor B. Kazantsev, and Sergey V. Stasenko. Artificial neural network model with astrocyte-driven short-term memory. *Biomimetics*, 8(5), 2023. ISSN 2313-7673. doi: 10.3390/biomimetics8050422. URL `https://www.mdpi.com/2313-7673/8/5/422`.

[23] Leo Kozachkov, Jean-Jacques Slotine, and Dmitry Krotov. Neuron-astrocyte associative memory. *arXiv preprint arXiv:2311.08135*, 2023.

[24] Leo Kozachkov, Ksenia V Kastanenka, and Dmitry Krotov. Building transformers from neurons and astrocytes. *Proceedings of the National Academy of Sciences*, 120(34):e2219150120, 2023.

[25] Samuel Cortinhas. Cats and dogs image classification. https://www.kaggle.com/datasets/samuelcortinhas/cats-and-dogs-image-classification, 2022.

[26] Ahmed Gamaleldin, Ahmed Atef, Heba Saker, and Ahmed Shaheen. Fire dataset. `https://www.kaggle.com/datasets/phylake1337/fire-dataset/data`, 2018.

[27] Sachin Panchal, Ankita Naik, Manesh Kokare, Samiksha Pachade, Rushikesh Naigaonkar, Prerana Phadnis, and Archana Bhange. Retinal Fundus Multi-Disease Image Dataset (RFMiD) 2.0, January 2023. URL `https://doi.org/10.5281/zenodo.7505822`.

[28] Jiancheng Yang, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister, and Bingbing Ni. Medmnist v2-a large-scale lightweight benchmark for 2d and 3d biomedical image classification. *Scientific Data*, 10(1):41, 2023.

[29] Ruhan Liu, Xiangning Wang, Qiang Wu, Ling Dai, Xi Fang, Tao Yan, Jaemin Son, Shiqi Tang, Jiang Li, Zijian Gao, Adrian Galdran, J.M. Poorneshwaran, Hao Liu, Jie Wang, Yerui Chen, Prasanna Porwal, Gavin Siew Wei Tan, Xiaokang Yang, Chao Dai, Haitao Song, Mingang Chen, Huating Li, Weiping Jia, Dinggang Shen, Bin Sheng, and Ping Zhang. Deepdrid: Diabetic retinopathy—grading and image quality estimation challenge. *Patterns*, page 100512, 2022. ISSN 2666-3899. doi: https://doi.org/10.1016/j.patter.2022.100512. URL `https://www.sciencedirect.com/science/article/pii/S2666389922001040`.

[30] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal

Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

[31] Yuliya Tsybina, Susan Gordleeva, Mikhail Krivonosov, Innokentiy Kastalskiy, Alexey Zaikin, and Alexander Gorban. Modelling working memory in neuron-astrocyte network. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6, 2021. doi: 10.1109/IJCNN52387.2021.9533307.

[32] Yuliya Tsybina, Innokentiy Kastalskiy, Mikhail Krivonosov, Alexey Zaikin, Victor Kazantsev, Alexander Gorban, and Susanna Gordleeva. Astrocytes mediate analogous memory in a multi-layer neuron-astrocytic network, 2021. URL `https://arxiv.org/abs/2108.13414`.

[33] Morten W Fagerland and Leiv Sandvik. The Wilcoxon-Mann-Whitney test under scrutiny. *Stat. Med.*, 28(10):1487–1497, May 2009.

[34] Verónica Bolón-Canedo, Laura Morán-Fernández, Brais Cancela, and Amparo Alonso-Betanzos. A review of green artificial intelligence: Towards a more sustainable future. *Neurocomputing*, 599:128096, 2024. ISSN 0925-2312. doi: https://doi.org/10.1016/j.neucom.2024.128096. URL `https://www.sciencedirect.com/science/article/pii/S0925231224008671`.

[35] João Filipe Oliveira, Vanessa Morais Sardinha, Sónia Guerra-Gomes, Alfonso Araque, and Nuno Sousa. Do stars govern our actions? astrocyte involvement in rodent behavior. *Trends Neurosci.*, 38(9):535–549, September 2015.

[36] Ciaran Murphy-Royal, Shinung Ching, and Thomas Papouin. A conceptual framework for astrocyte function. *Nat. Neurosci.*, 26(11):1848–1856, November 2023.

[37] Ilya A Zimin, Victor B Kazantsev, and Sergey V Stasenko. Artificial neural network model with astrocyte-driven short-term memory. *Biomimetics*, 8(5):422, 2023.

[38] James M Robertson. The astrocentric hypothesis: proposed role of astrocytes in consciousness and memory formation. *J. Physiol. Paris*, 96(3-4):251–255, April 2002.