Utrecht University

MSc Computing Science

Track: Algorithmic Data Analysis

Master thesis

Attention-guided Visual Learning, A Computational Model

by

Maaike Koninx

6303919

March 23, 2020

25 EC

Daily Supervisor:
Prof. Sander Bohté

Examiners:
Dr. Ronald Poppe
Dr. Dirk Thierens

Abstract

Our behavior is driven by a small subset of all information available to us. As our processing resources are limited, we select information to attend to in order to learn from our environment. Visual attention has been studied for many decades. Still, current attentional models do not explain how attentional modulations affect trial-and-error learning in the visual cortex. This study is the first to define synaptic plasticity as a function of attentional modulations observed prior to receiving rewards. The attention-modulated Hebbian plasticity rule is used to simulate attention-guided learning for a series of classification tasks. Despite exclusively receiving reward feedback for the predicted label, our attention-guided reinforcement learning framework is able to perform comparably to supervised error-backpropagation. This holds for datasets with up to 100 class labels. Our results are obtained by redefining learning to reflect biological mechanisms which ultimately govern behaviour.

Contents

1 Introduction

Our behavior is driven by a small subset of all information available to us. As our processing resources are limited, we are required to select information to attend to in order to learn from our environment. Attention is therefore considered a selection process [16], as well as an increased bias towards relevant information encoded in neural populations [6]. Neurophysiological studies have focused on measuring and describing this process in the visual cortex for many decades. Despite a large body of literature on visual attention, we lack a solid understanding of the mechanisms behind attention-guided visual learning. Computational models that simulate complex interactions in the brain can be used to implement attentional mechanisms by biasing network activity to relevant input [12]. However, current attentional models either do not reflect biological learning mechanisms in the visual cortex [11] [18] [2] [32] [5], do not account for attention as part of the learning process [13] or fail to explain how learning relates to known attentional modulations [26] [23].

Individuals generally learn rewarding behaviors from trial-and-error interactions with their environment. This *reinforcement learning* is guided by reward prediction errors [25], which are believed to be propagated by the dopaminergic system to neurons that contributed towards the reward-associated behavior. Here, prediction errors are believed to alter the synaptic strength between neurons to either enforce or discourage the associated behavior [25] [26]. Hebbian plasticity states that synaptic strength reflects the causality between presynaptic and postsynaptic firing [26]. This well-established belief is more commonly stated as "cells that fire together, wire together". Moreover, neurophysiological studies find that attention alters neural firing rates prior to receiving a reward [16] and increases performance accuracy and speed [6]. It can therefore be suggested that learning mechanisms are guided by these attentional firing modulations.

How do attentional mechanisms aid learning? This study addresses this question by describing a biologically plausible learning framework that simulates attentional firing modulations and integrates these into trial-and-error learning. The attention-modulated Hebbian plasticity rule is used to simulate attention-guided learning in the visual cortex for a series of classification tasks. Despite exclusively receiving reward feedback for the predicted label, our attention-guided reinforcement learning framework is able to perform comparably to supervised error-backpropagation, which provides feedback for all class labels. This holds for datasets with up to 100 class labels, i.e. where there is 100 times more feedback information for supervised learning than for trial-and-error learning. Furthermore, while attention generally increased the performance of our learning rule, our results did not indicate significant attentional performance gains. However, we propose a slight addition to the learning framework that could result in attentional performance gains that may outperform error-backpropagation. The learning framework proposed in this study is the first to define synaptic plasticity as a function of attentional modulations observed prior to receiving rewards.

It is important to mention that while this study aims to simulate attentional processes, our goal is not to achieve biological 'correctness'. For instance, we do not implement a spiking network which would be more coherent with biological neural networks. Moreover, current artificial neural networks do not account for types of neurotransmitters and connections that are either excitatory or inhibitory, i.e. that cannot change sign. Many aspects of implementing these processes are still unknown and beyond the scope of this study.

Before discussing the proposed learning framework it is important to review previous attentional research. We review previous work on attentional modulations of behavior in section 2.1, followed by attentional modulations of neural populations in section 2.2, and attentional models in section 2.3. We discuss reinforcement learning and current biologically plausible learning frameworks in section 2.4. We argue that current work calls for the attention-guided learning framework from section 3. Results from the experiments from section 4 are detailed in section 5. We conclude our study in section 6.

2 Background

2.1 Attending to stimuli

Before we discuss different attentional modulations, it is important to differentiate between types of attention. Attention is guided by top-down or bottom-up information or both [4]. Top-down information like intrinsic goals or knowledge is propagated from higher to lower cortical areas, while bottom-up information like the 'saliency' or 'unexpectancy' of stimuli flows from lower to higher areas. Tasks that require a visual search for a unique conjunction of features between several potential targets require both top-down and bottom-up information. The studies we discuss in the following sections focus on top-down and bottom-up attentional modulations as a result of contextual information. In contrast, attentional modulations due to non-contextual 'alarming' stimuli are located in more ventral regions that likely serve as an attentional circuit-breaker on all sensory modalities [7].

Another important distinction to be made is between types of stimuli that drive attention. Attention can be directed to a location in the visual field, features that correspond to specific aspects of a stimuli such as color, or objects. The mechanisms behind these attention types are most likely not separate entities. As we will explain in the next section, spatial attention is likely a special form of feature-based attention, which likely represents the same mechanism as object-based attention [22]. While feature-based attention is most profoundly measured in the visual cortex [4], object-based attention is measured in higher cortical areas which are generally specialized towards categories such as faces, body parts and scenes [22]. Spatial attention is measured in the visual system, as well as the frontal, parietal and midbrain areas that control eye movement [4].

Attentional modulations can be tested through spatial or feature cueing before the target-stimulus onset, see Figure 1. Spatial cueing directs attention to a location in the visual field while feature-based cueing directs attention to a target direction, colour or motion. Target cueing increases performance accuracy as well as reaction speed in many different types of visual tasks [6]. However, these *attentional performance gains* decrease with the level of uncertainty. That is, when the size of the attended area or range of potential target-stimuli increases in spatial and feature-based attentional tasks respectively. This gain-deterioration is most evident in hard tasks, such as low-contrast target detection tasks. Similarly, attentional gains increase while performance accuracy deteriorates with the number of distractors, even when these are known to be behaviorally irrelevant [6].

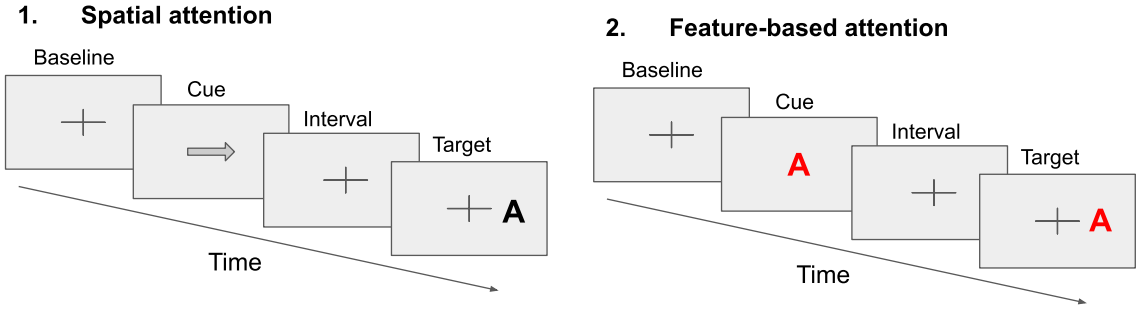

Figure 1: Detection tasks. *Left*: a spatial cue directs attention to the right visual field before target-stimulus onset. *Right*: a feature cue directs attention to a red letter 'A'. The target-stimulus can appear on both sides of the fixation point. Attentional modulations are measured during the cue-target interval.

While some studies reported attentional performance gains when subjects were exposed to high levels of external noise [9], other studies found similar attentional gains in zero-noise conditions [14]. This suggests that attention enhances the target-stimulus response as well as suppresses noise-stimuli. Interestingly, many studies report increased perceived target contrast and resolution in cueing tasks compared to non-cueing tasks [6]. This indicates that attention alters our perception of the world around us according to what we attend. Hence, attention 'highlights' what we deem relevant or important at the expense of a true representation.

2.2 Neurophysiology of attentional modulations

Attention in the visual stream This study will focus on modelling attention in the *visual system*, see Figure 2. This system receives feedforward information from the retina in specialized areas of the fore- and midbrain [4]. From here, information is passed along the ventral and the dorsal pathway of the visual cortex, and on to higher visual association areas. Neural modulations as a result of both spatial and feature-based attention generally increase with this feedforward flow. We will discuss these modulations in the next section. Other areas that exhibit attentional modulations are the frontal eye fields (FEF) which controls eye movements, the

lateral intraparietal cortex (LIP) which receives information from both streams of the visual cortex, and the prefrontal cortex (PFC) which is involved in intrinsic goals, planning and decision making [4].

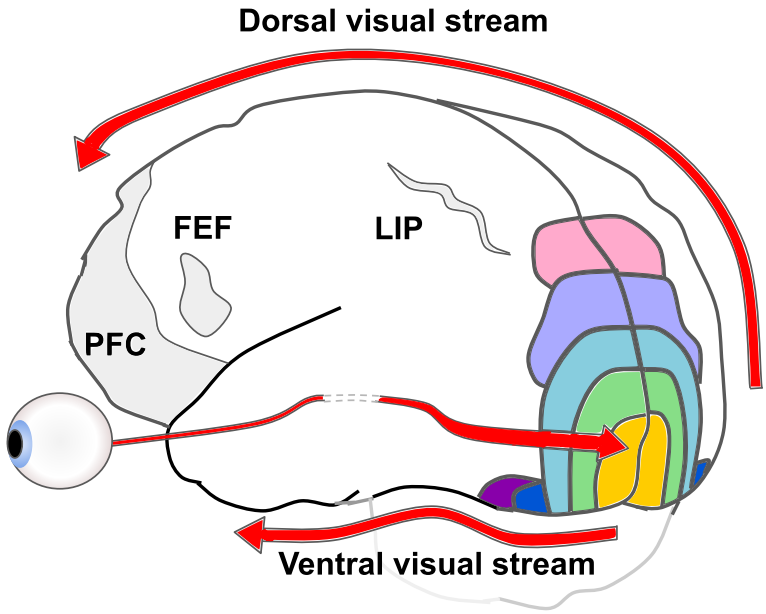

Figure 2: Illustration of the human visual system. Information from the retina is passed along the dorsal and ventral streams of the visual cortex. Other areas include FEF, LIP and PFC.

Top-down information from the frontal cortex to lower sensory areas is either communicated by back-propagating through the visual stream or through alternate connections to lower areas. Such 'shortcuts' can be found from FEF and the prefrontal cortex to visual cortices [4]. These pathways are believed to inform lower sensory areas of spatial and feature-based biases. Moreover, top-down information encoding intrinsic knowledge and goals and bottom-up information encoding the saliency of stimuli are combined to construct *priority maps* of visual space [4]. These maps indicate behaviorally important areas in the visual field and govern eye movements accordingly. Priority maps are likely located outside of the visual system, e.g. in FEF and parts of the midbrain.

While priority maps do not contain any feature-based information, neurons in the visual system are often highly specialized in encoding features [4]. Hence, mechanisms that govern eye movement and feature-based attention have separate neural substrates and likely interact bi-directionally through connections in the forebrain. As we will explain in section 2.3, hard attentional computational models produce sequences of relevant locations in images from which cropped areas are fed into the network, similar to how priority maps induce human eye saccades over an image [1]. Since this study focuses on simulating attention in the visual cortex, which is not involved in governing eye saccades, hard attention will not be implemented.

Neural attentional modulations Information that feeds into neurons of the visual cortex generally encodes areas of the visual field known as the neuron's receptive field. The size of neuronal receptive fields increases with the visual feedforward flow as each area integrates information from the previous area, similar to deep convolutional networks [6]. Additionally, neurons along higher cortical areas exhibit selective firing responses to certain features such as a 'preferred' direction [17]. Neuronal feature-selectivity can be visualized by mapping the firing response as a function over a continuous feature range. A sharpened, i.e. narrowed, *tuning curve* is believed to indicate an increased ability to encode a particular feature [15].

When attention is directed to a stimulus in their receptive field, neurons along the visual stream are known to fire more, faster and more consistently following the stimulus onset compared to non-attended tasks [16]. These *attentional response gains* generally increase with the feedforward flow along the visual system and with task difficulty, as cueing increases the amount of relevant information encoded in neural populations most prominently in hard tasks. Additionally to altered response rates, delay time between stimulus onset and neural responses decreases with the size of attentional gains [16]. Furthermore, baseline activity increases prior to stimulus onset in the neural population encoding the attended area and decreases in neurons encoding non-attended areas [6].

A well-known study by True & Martinez-Trujillo (1999) [17] was the first to separate feature-based attentional modulations from spatial biases. Treue & Martinez-Trujillo (1999) excluded spatial noise by directing attention to a moving-dot pattern outside of the receptive field while maintaining the same pattern in the receptive field. Remarkably, they found that neural responses are multiplicatively scaled compared to directing attention inside the receptive field, see Figure 3. These results suggest that feature-based attention is spatially-invariant, i.e. alters response rates across the neural population. Moreover, Treue & Martinez-Trujillo (1999) argue that neural responses are multiplied by a set rate that represents the similarity between attended and preferred features. This *feature-similarity gain model* predicts increased response rates when neurons prefer the attended feature and a maximized response when this feature is observed in the receptive field.

While [17] found that attention multiplicatively scales tuning curves without changing their shape, see Figure 4a, other studies report altered tuning behavior. Some studies report sharpened, i.e. narrowed, tuning curves on individual or population levels [15], indicating attention increases selectivity. Yet others studied attentional modulations on tuning curves over a range of contrasts and found a horizontal shift towards low-contrast stimuli, indicating attention increases sensitivity, see Figure 4b. These results do not necessarily contradict one another. In a subsequent study [15], Treue & Martinez-Trujillo (2004) found that while attention multiplicatively scaled responses for individual neurons, sharpened tuning curves were found at the population-level. As we will see in section 2.2, a simple normalization model explains both multiplicative scaling, i.e. *response gain*, and increased contrast sensitivity, i.e. *contrast gain*.

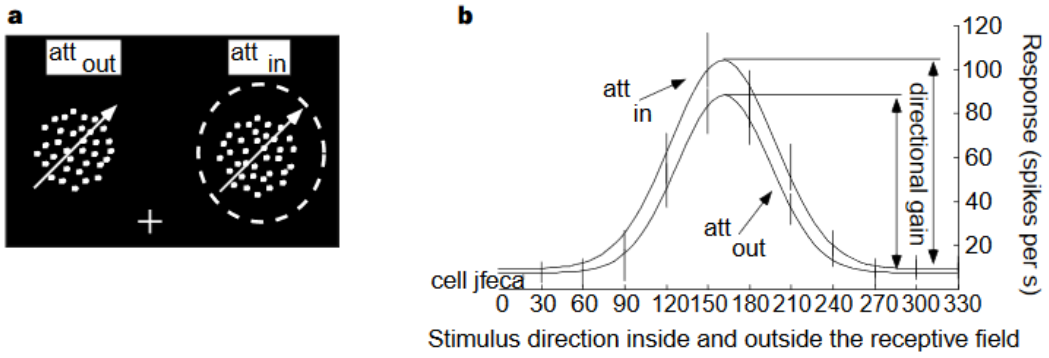

Figure 3: Attentional scaling of the direction-tuning curve, image from [17]. **(a)**: Attention was directed inside or outside the receptive field. **(b)**: Response rates were multiplicatively scaled.

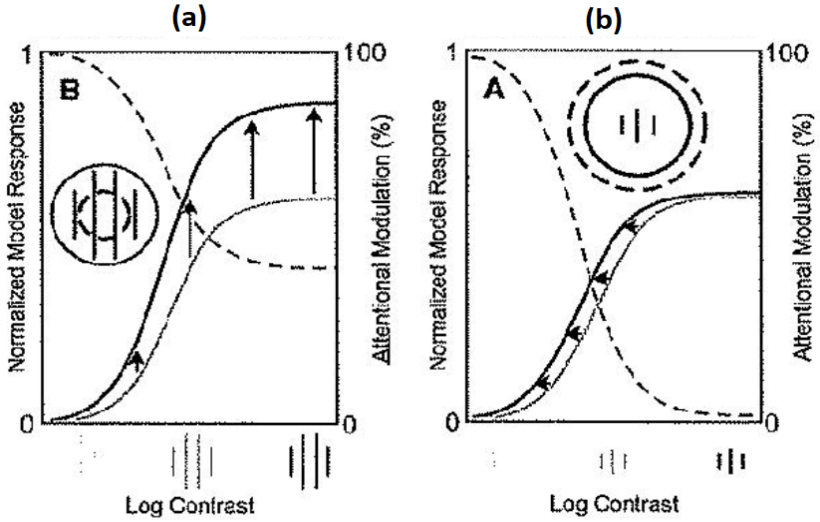

Figure 4: Attentional modulations of contrast-tuning functions, image from [24]. **(a)**: Response gain. **(b)**: Contrast gain.

Subsequent studies that isolated feature-based attentional modulations from spatial biases found many identical results to previous work. Feature-based attention increases response rates of neurons that prefer the attended feature, similar to spatial attentional gains to a stimuli in the receptive field [6]. Similar to spatial attention, feature-based attention is associated with decreases in pairwise variability correlations between similarly tuned neurons and increases synchronized high-frequency oscillations [16]. As a result, spatial and feature-based attention are argued to represent the same mechanism. Particularly, spatial attention may be a form of feature-based attention where neural responses are tuned to the receptive field. However, some studies report differences in attentional modulations [21], some of which we discuss in the next section.

To conclude the experimental results discussed in this section, we discuss attentional modulations related to object-based attention. Attentional response gains are high-

est in areas that selectively respond to the target-object, such as the fusiform face area for attending faces [3], and these modulations are spatially-invariant. Moreover, object-based attention increases firing patterns related to the target-category while decreasing distractor-related patterns [29] and is associated with increased gamma-oscillations [3]. These results suggest object-based attention may represent the same mechanism as feature-based attention in higher cortical areas.

Models for attentional modulations There are many theories that attempt to unify experimental results. Attention has been argued to shrink receptive fields [19] or bias response rates [8] of neurons that encode the attended stimulus. While Treue & Martinez-Trujillo (1999) explain their results with the feature-similarity gain model, they can alternatively be explained with the feature-matching hypothesis [15]. This model states that attentional gains represent the similarity between the attended and perceived features irrespective of tuning. However, a subsequent study [15] failed to find a correlation between attentional gains and the similarity between the attended and perceived stimulus. Lastly, attention has been argued to increase contrast sensitivity.

While these theories explain part of the experimental results, none explain why some studies observe contrast gains while others find response gains of the contrast-tuning function. Furthermore, none explain why switching attention between a preferred and its opposite, i.e. anti-preferred, feature in a neuron's receptive field alters its firing rate [17]. Hence, a study that was able to explain all these results with a simple theoretical model [24] provided a significant breakthrough in attentional research. This *normalization model* by Heeger & Reynolds (2009) [24] states that attention acts as a response multiplier to the attended stimulus, but the resulting neuronal activity is normalized over the population response according to similarity in receptive field location and tuning.

For example, when a single stimulus is presented in the receptive field, the normalization model states that the size of the stimulus determines the normalization rate. If the stimulus is small, few other neurons will perceive the stimulus and normalization only scales down large population responses, i.e. responses to high contrasts, resulting in contrast gain. If the stimulus is big, large cumulative population responses will be scaled down by normalization even for low contrasts, resulting in response gain. Furthermore, when attention is directed to either of two opposite stimuli in the receptive field, the normalization model claims that the neuron's response will be normalized over the activity from adjacent neurons that are tuned to the attended stimulus. When attention is directed to the preferred feature, the neuron's attention-increased response will be little affected by this population's activity. However, when attention is directed to the anti-preferred feature, the neuron's response drops significantly due to cumulative activity of the population that prefers this feature. This theoretical model is the first to explain the modulations observed when switching attention between stimuli in the receptive field.

The non-linearity of the normalization model offers some interesting methods to

identify the source of attentional modulations of the feed-forward flow. A recent study [28] micro-stimulated lower cortical neurons to show that cross-area attentional modulations are best explained by altered connections to higher areas, as opposed to local neural response modulations. However, some studies argue that inter-neuronal differences in attentional gains, are better explained by a *tuned normalization model*. This model states that normalization scales the sum of the neuron's tuned responses L_i over all stimuli i with contrast c_i in the receptive field. Moreover, this normalization strength is a weighted sum of population responses to certain stimuli or locations i where weights α_i represent the neurons receptiveness to i. For instance, intrinsic weights in a spatially-tuned neuron represent its receptiveness to population responses from certain areas in the receptive field. In contrast, the normalization model from [24] assumes this responsiveness is a smooth function over the similarity to other neuron's receptive fields and tuning. In this model, attention β_A acts as a rate multiplier to the attended stimuli A both for the individual response and the normalization strength, see Table 1.

Model	Attention type	Neural response rate
Tuned normalization	Any	$\frac{\beta_A c_A L_A + c_N L_N}{\beta_A \alpha_A c_A + \alpha_N c_N + \sigma}$
EMS-tuned normalization	Spatial	$\frac{\beta_A c_A L_A}{\beta_A c_A + \alpha_N c_N + \sigma} + \frac{c_N L_N}{\beta_A \alpha_A c_A + c_N + \sigma}$
	Feature-based	$\frac{\beta_A c_A L_A}{c_A + \alpha_N c_N + \sigma} + \frac{c_N L_N}{\alpha_A c_A + c_N + \sigma}$

Table 1: Several attentional models for neural response rates to two stimuli A, N when attending A. Here, β_A signifies the attentional strength, and c_i, L_i and α_i the contrast, individual tuning and individual weighting of $i \in \{A, N\}$ respectively. Lastly, σ is a constant.

Ni & Maunsell (2017) [20] found that while the spatially and stimulus-tuned normalization models explained a majority of neural responses in spatial attentional tasks, it failed to capture some neural response variability when multiple stimuli were present in the receptive field. Instead, neural responses could be approximated as a sum of normalized responses, rather than a normalized cumulative response, over all stimuli in the receptive field. When spatially tuned, this EMS-tuned normalization model explained up to 92% of their data. However, a subsequent study [21] found that feature-based attentional modulations were not correlated to normalization strength, indicating an essential difference to spatial attention. Instead, Ni & Maunsell (2019) [21] argue that feature-based attention acts as a rate multiplier to the attended feature in the numerator but not the denominator, i.e. manipulates response rates prior to normalization but not normalization strength. Hence, they argue that a single attentional signal is shared by both types of attention but this signal alters normalization differently.

2.3 Computational attentional models

Convolutional networks are inspired by the human visual system [13]. Artificial neurons, similar to biological neurons, have receptive fields from which they receive information through weighted connections. Moreover, deep convolutional networks (DNNs) have many layers which accumulate specialized information over larger areas in higher layers, similar to the visual system. Hence, DNNs provide a ready framework for implementing and validating the theoretical attentional models discussed in the previous section. Previous computational studies have focused on implementing object segmentation frameworks, such as Mask R-CNN [11]. However, these models are not designed to simulate attentional mechanisms for classification tasks, are computationally expensive and scale poorly with respect to input size. Hence, we will discuss alternative computational models of visual attention.

One important distinction to be made is whether computational attentional models have differentiable loss functions with respect to the input. Computational models that implement *soft attention* have loss functions that can be expressed as a differentiable function of the network weights and can be trained through back-propagation in a supervised learning setting. On the other hand, models that implement *hard attention* simulate human eye saccades by producing a location sequence in images from which cropped areas are used for classification [1]. Hence, predictions made by these models have no known 'correctness', but instead yield rewards that can be trained through reinforcement learning. Training such models requires no labelling of 'correct' saccades if we know whether a location sequence results in a correct classification, i.e. a reward.

Hard attention was first implemented in [18] to bypass computational costs associated with large images, as computational complexity in CNNs scales linearly with the number of pixels. At each time step of a recurrent neural network (RNN), frames of variable resolutions are cropped from the input image following the predicted location from the previous state and fed into the current state, which predicts the next location. While this recurrent attention model predicts a single object category after a fixed number of saccades, a deeper alternative [1], repeats this procedure until a stop-sign label is predicted in order to detect multiple objects. Additionally, this deep recurrent attentional model (DRAM) performs classification in a prior recurrent layer to prevent contextual predictions from biasing classifications, see Figure 2.3a. This model outperforms many regular CNNs on large images as its computational complexity can be controlled independently of input size by manipulating the cropped frame resolution.

Recall that attentional mechanisms that allow for feature-based biases are measured in different neural substrates than mechanisms that govern eye saccades [4]. Therefore, biologically plausible computational models for feature-based attention differ from models that implement hard attention. Models for feature-based attention create a bias towards relevant features based on their contribution towards the correct prediction. As such models are few and very complex, we will first discuss their

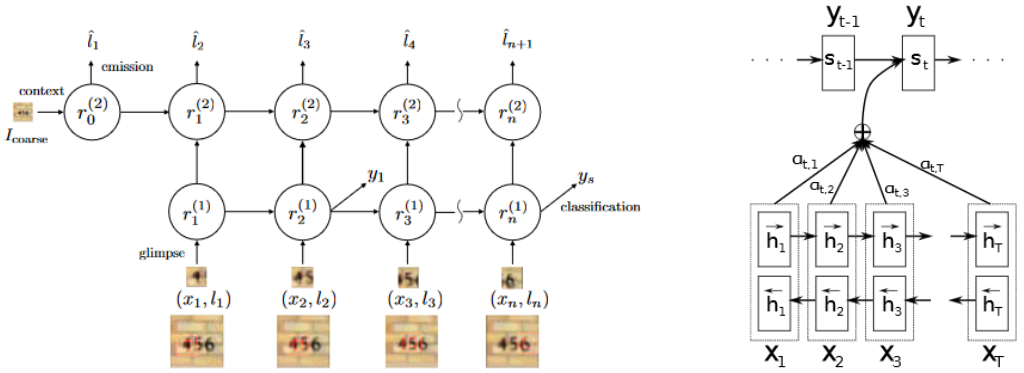

Figure 5: Two computational models implementing attention. **(a):** DRAM implements hard attention, image from [1]. **(b):** An encoder-decoder model implements soft attention, image from [2].

origin from language models that implement soft attention.

Soft attention in language models is the mechanism of 'highlighting' relevant words through customized weighting of an input sentence for each prediction [2]. Specifically, these models consist of two recurrent networks, an *encoder* that embeds input words into a vector representation and a *decoder* that applies state-dependent weighting of each embedded word for each prediction step, see Figure 2.3b. The encoder often consists of stacked RNNs that perform forward $\overrightarrow{h_i}$ and backward passes $\overleftarrow{h_i}$ and outputs a concatenated embedding $h_i = [\overrightarrow{h_i}, \overleftarrow{h_i}]$ for each word i, called an annotation. This ensures that each annotation contains information on its preceding and succeeding words [2]. At each timestep t, the decoder of state s_t computes a weighted function $c_t = \sum_i a(s_{t-1}, h_i)h_i$ over this embedded sentence to find the prediction y_t that maximizes the conditional probability $P(y_t|y_{t-1}, s_t, c_t)$. While many deviations from this base architecture exist, the key component to these soft attentional models is the customized weighting of complete input sentences for each prediction. This allows RNNs to considerably increase their ability to capture long-distance dependencies.

Soft attentional models yield high performances on different language tasks such as language translation [2] and query-answering [12]. However, recurrent networks are inherently non-parallelizable as each state depends on its previous state, and are therefore not fit to train on large sequences. Some models attempt to solve this by embedding sequences in a two-dimensional space and training convolutional networks to 'soft-attend', i.e. increase weighting from, certain areas of this input [10]. By switching to these convolutional networks however, soft attentional models encounter performance loss on large sequences as CNNs are localized due to limited receptive fields. Hence, these models offer no sufficient solution for sequences with long-range dependencies.

Transformers [32] apply a form of soft attention in parallel self-attentional layers called blocks which are considerably more parallelizable and less localized than recurrent or convolutional models. Each self-attention block produces key (K), value (V) and query (Q) matrices as weighted functions of embedded words. As each matrix is designed to encode some type of information on each word, this information can be combined through highly-parallelizable matrix multiplications. That is, the t-th softmaxed row of $Q \cdot K^T$ resembles the relative soft-attentional weights of the sequence for output t, and its product with V equals a weighted function, similar to c_t in the decoder, of the input sentence for that output. After each self-attention layer, the outputs from all attention blocks are concatenated for each timestep, combined with the sublayer through a residual connection, normalized and fed into a feedforward layer. As the details on these operations and the transformer architecture exceeds the scope of this study, we refer the interested reader to [32].

Many computational models alter or extend this transformer structure to achieve state-of-the-art performance on language modelling benchmarks. For instance, [31] replaces each self-attentional layer and its subsequent feedforward layer by an all-attention layer that aggregates embedded input vectors and non-contextual vectors encoding general task-relevant information. While a positional encoding in each embedded word allows the model to apply attention over a one-dimensional sequence, encoding such information from two-dimensional data, such as an image, is not straightforward. Simply extending positional encodings to two dimensions violates translation equivalence, i.e. alters output in case of an input shift, [5] and likely does not benefit performance. However, a recent study [5] applies relative 2-dimensional positional encodings to solve this. Furthermore, Bello et al. (2019) find that this relative self-attention mechanism performs best when concatenated with feature maps in a convolutional network.

While self-attentional models such as [5] achieve promising performances on image classification datasets, no evidence exists that they account for human attentional mechanisms. Self-attention requires multiplication between variables while, to our best knowledge, neurons of the visual system cannot multiplicatively combine their output. Whether indirect neural mechanisms can account for soft-attention, we will leave for future research. While attentional modulations in the visual system cannot be explained by any previously discussed computational models, their implementations can be fairly simple. The feature-similarity gain model states that attention multiplicatively scales neural responses according to the similarity between their feature-tuning and the attended stimulus. Hence, a simple implementation of this model would compute neuronal tuning values and multiplicatively increase neural output according to how much it corresponds with the attended feature.

A recent study [13] was the first to attempt this on a population level. That is, Lindsay & Miller (2018) multiplicatively scaled the output of feature maps of a pre-trained DNN according to their tuning value for the attended category. These tuning values were computed from the normalized feature map output as a function over the object-categories. That is, if $r_f(c)$ is the population-averaged output of

feature map f to category c, then the tuning value $t_f(c)$ of f equals its normalized activity over all N categories, i.e.

$$t_f(c) = \frac{r_f(c) - \bar{r_f}}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}(r_f(j) - \bar{r_f})^2}}, \quad \text{where} \quad \bar{r_f} = \frac{1}{N}\sum_{k=1}^{N} r_f(k)$$

Hence, t_f is a discrete tuning curve of the feature-map population and determines how attention to an object-category multiplicatively scales the feature map output. That is, attention of strength β to category c scales the output of neuron x^f in feature map f to

$$x^f_{Att}(c) = (1 + \beta t_f(c))x^f \tag{1}$$

While this describes *how* attention is implemented for a given category, it does not yet present a model that can dynamically apply attention to increase image classification accuracy. However, the output layer of the DNN consists of binary classifiers that detect the presence or absence of each category and separately apply attention to the feature maps accordingly. Hence, each binary classifier for category c applies c-specific tuning to each feature map activation and this process is executed for all classifiers independently. The resulting model can dynamically apply attention to detect which categories are present in an image.

Lindsay & Miller (2018) found that applying this tuned-attention model in all layers of their DNN increased image classification performance. This indicates that the category-specific scaling of neural population activity aids in detecting each category separately, which leads to a higher overall accuracy. Remarkably, this attentional performance gain maximizes when applying attention solely in the final layer, i.e. to 18.8% and 22.8% in merged images and 2x2 image grids respectively. This performance gain decreased exponentially when solely applying attention in earlier layers. While Lindsay & Miller (2018) argue that the network fails to propagate attention, their results show that tuning quality, i.e. the maximum tuning curve value, is considerably lower for earlier layers. Hence, attention at lower layers induces little selectivity, suggesting that this study is limited to simulating top-down attention. The increase in tuning quality in higher layers corresponds to the increased sharpening of tuning curves measured in higher cortical areas [4]. However, similar tuning behavior and performance gains were found when replacing $t_f(c)$ with the average prediction-error gradient over feature map f for category c in Equation 1. Hence, attention might alter population responses according to tuning as well as to their cumulative prediction error.

2.4 Attention & Reward

While Lindsay & Miller (2018) [13] provide a simple and elegant method for modelling attention in the visual system, their model does not account for attention as part of the learning process. That is, their model applies attention to a pre-trained network where no more learning occurs. In reality, attention is more likely to aid learning [25]. Moreover, this learning process in reality likely does not resemble supervised learning as labelled data defining correct and incorrect output-labels is

rarely available to us. Individuals generally learn rewarding behaviors from interacting with their environment. This is called reinforcement learning. As many learning mechanisms in the brain rely on reinforcement learning, it is crucial for a computational model on attention-guided learning to enforce a biologically plausible reinforcement learning rule.

Models that apply reinforcement learning provide solutions to tasks for which no ground-truth answer exists, e.g. because it is not known or the aim is to outperform known strategies. Hence, learning is not restricted to optimizing classification accuracy on training data, as is the case with the supervised form of learning, but also involves exploration of an unknown environment. It therefore imposes a trade-off between choosing rewarding behavior or pursuing higher rewards in an unknown territory. Furthermore, there are different methods on how expected rewards determine behavior in a current state s based on previous behavior in s. For instance, some methods simply choose the behavior that maximizes the average received reward over past behavior in s while others choose to maximize the predicted reward based on reward predictions of succeeding states. The latter method is called temporal difference learning [27].

The neural mechanisms behind reinforcement learning have long been subject to neurophysiology studies. Temporal difference learning in particular has received much attention as its reward predictions correlate strongly to neural responses in the dopaminergic system, which is closely linked to reward expectation [27]. Differences between received and expected rewards, or reward prediction errors, are believed to drive reinforcement learning by altering synaptic strength between neurons [25], similar to artificial networks. When receiving (or not receiving) a reward, information encoding reward prediction errors is propagated by neuromodulatory systems such as the dopaminergic system, to the neural population that formed the prediction. However, a main issue with this system is credit-assignment, i.e. which connections contribute to the prediction and should be altered. A number of neurophysiology studies suggest that synapses that contribute to a prediction receive an attentional feedback signal from higher cortical areas which marks them with a biochemical tag [25]. These tags form eligibility traces that gate which synapses are altered when receiving reward feedback signals.

As we can only learn what we attend, attention can be considered as a gating mechanism for learning. This mechanism of attentional gating of *which* synapses should be altered and sending a feedback signal determining *how* they should be altered, should be accounted for when modelling attention-guided learning. The attention-gated reinforcement learning scheme AGREL [26] defines a reward system that accounts for this attentional gating and that backpropagates reward prediction error gradients through separate connections for networks consisting of a hidden layer Y and a softmax layer Z. Each backward connection $w'_{j,i}$ from neuron j to i determines the plasticity of the forward connection $w_{i,j}$ and is updated according to the same plasticity rule so that it approximates $w_{i,j}$ during training. This plasticity

rule states that weights are updated according to

$$\Delta w_{i,j} = \alpha R_i f b_j(\delta) \tag{2}$$

where R_i is the activity of presynaptic neuron i, α is the learning rate and $fb_j(\delta)$ is the feedback received in postsynaptic neuron j about the prediction-error δ. Each output $k \in Z$ is predicted with probability R_k, after which the value of the predicted output is set to 1 while others are set to 0. The prediction-error δ for predicting output unit s is set to $(1-R_s)/R_s$ if the prediction is correct and -1 for an incorrect prediction. The feedback on this prediction-error that is backpropagated through postsynaptic neuron j equals

$$fb_j(\delta) = \begin{cases} \delta\, R_j & \text{if } j \in Z \\ \delta\, R_j\, w'_{s,j}\, (1 - R_j) & \text{if } j \in Y. \end{cases}$$

Note that as $R_k = 0$ for all unpredicted outputs $k \in Z$, there is no feedback to weights that did not contribute to the prediction. This simulates the attentional gating mechanism. The dopaminergic feedback machanism is simulated by fb_j and the backward connections $w'_{i,j}$.

While AGREL provides biologically plausible learning rules for networks up to two layers, the cortex consists of many neural layers. Hence, for a biologically plausible model for attention-guided learning, the AGREL rules need to be generalized to fit deeper networks. The Q-AGREL [23] rule states that for a network with any number of hidden ReLU layers $Y^{(1)}, \ldots, Y^{(N)}$ and a linear output layer O, the feedback $fb_j^{(n)}(\delta)$ received from node $j \in Y^{(n)}$ from layer $n \leq N$ equals

$$fb_j^{(n)}(\delta) = g'(j) \sum_{y \in Y^{(n+1)}} w'_{y,j} fb_y^{(n+1)}(\delta), \quad \text{where} \quad g'(j) = \mathbb{1}_{\{j>0\}}.$$

Moreover, $fb_j^{(O)}(\delta) = \delta g'(j)$, where the prediction-error δ is determined by the mean squared error loss function and equals $\delta = 1 - R_s$ for correct predictions and $\delta = -R_s$ otherwise. The function g equals the gradient of the activation function, which is assumed to be the ReLU function but can be substituted for other functions. By using the ReLU function, g ensures only the synapses to nodes that contributed to the prediction, i.e. with activity greater than 0, are altered. The resulting Q-AGREL framework is similar to error-backpropagation of the loss over the predicted output node. Moreover, Pozzi, Bohté & Roelfsema (2019) [23] have shown that Q-AGREL performs comparably to supervised learning in a range of image classification tasks when trained on 1.5-2.5 times the number of epochs.

While Q-AGREL provides a biologically plausible model of prediction-error backpropagation through the dopaminergic system, it contains some assumptions that likely do not comply with biological learning processes. Most importantly, attention in Q-AGREL solely acts as a gating mechanism on the backpropagation of reward predictions by dopaminergic neurons that alter the synapses of feedforward neurons directly. However, neurophysiology studies on attentional response gain [17]

[15] suggest that attention-guided learning may be a result of neural gain modification. That is, attention, guided by top-down signals, multiplicatively scales neural responses and synaptic strength is altered accordingly. In contrast to Q-AGREL, such a mechanism would be able to explain attentional gain observed in neural populations prior to receiving a reward and would therefore be more biologically plausible.

3 Attention-guided learning rules

In the previous section we argued that current attentional models either do not reflect biological learning mechanisms in the visual cortex or fail to explain how learning relates to known attentional modulations. Here we present a biologically plausible rule for attention-guided learning in the visual system. To our best knowledge, this is the first model that describes reinforcement learning as a result of attentional response gain observed in neurons of the visual system.

Attentional gain can be simulated by multiplicatively altering the activity of neuron y with attentional term β_y during training. For a network with layers $Y^1, \ldots, Y^N$, the neural activity of $y \in Y^n$ with activation function g therefore equals

$$y_{out} = g((1 + \beta_y) \cdot y_{in}), \quad \text{where} \quad y_{in} = \sum\nolimits_{x \in Y^{n-1}} w_{x,y}\, x_{out} + b_y. \tag{3}$$

Here, x_{out} signifies the neural output of the presynaptic neuron x and Y^0 signifies the network input. When using the ReLU activation function, the attentional term will simply scale the neural output of y with β_y, see Figure 6.

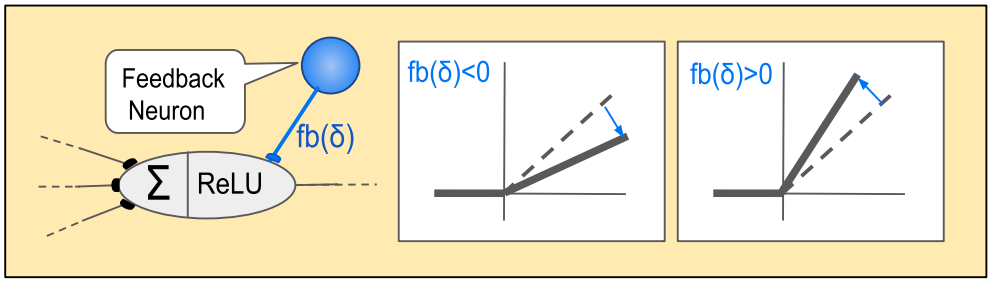

Figure 6: Illustration of how feedback $fb(\delta)$ on the prediction error δ alters the attentional gain for a ReLU-activated neuron.

Action-selection. For each training sample, all β's are initialized to 0, indicating no initial attentional modulation. The output of this unattended state determines the network prediction in the action-selection phase. As reinforcement learning imposes a trade-off between exploiting known rewards and exploration, the network will select the highest output unit in the output layer with probability p and explores another unit with probability $1 - p$. In this study, $p = 0.02$ unless stated otherwise. If $\sigma = \sigma_1, \ldots, \sigma_k$ signifies the softmax probabilities over the k output units, then

output unit s is selected according to

$$s = \begin{cases} k & \text{with probability } \sigma_k\,(1-p) \\ \operatorname{argmax}\sigma & \text{with probability } p. \end{cases}$$

After predicting unit s, the network will generate a target output $\epsilon_s = e_1, \ldots, e_k$ that reflects this prediction. Here, e_s equals 1 while other output units $e_{i \neq s}$ equal 0. Hence, ϵ_s equals the one-hot encoding of the selected class s.

Attentional phase. After selecting class s, the network is optimized to fit the predicted target output ϵ_s by altering the attentional gain using gradient descent. By optimizing the network to its prediction prior to receiving the reward prediction error, attention is allowed to alter the network activity prior to learning, similar to attentional modulations observed in the visual system. Here, a loss function that corresponds to cross-entropy loss is used to compute reward prediction errors as the relative output probabilities best reflect neural competition in higher cortical areas.

The loss between the output probabilities σ and predicted target output ϵ_s equals

$$Loss(\epsilon, \sigma) = \Sigma_i\, e_i \log \sigma_i,$$

the error δ for each output unit i equals $\delta_i = e_i - \sigma_i$. Hence, $\delta_i = 1 - \sigma_i$ for $i = s$ and $-\sigma_i$ otherwise. If $fb_y(\delta)$ denotes the feedback received by neuron y on prediction error δ, then a simple derivation of Equation (3) defines the update of attentional term β_y as

$$\Delta\beta_y = \alpha\, y_{in}\, fb_y(\delta) \tag{4}$$

where α is the attention rate. Here, attentional terms will be capped at -0.5 and 0.5, to allow a maximum attentional gain of 50% as this simulates physical limitations of response gain in the visual cortex [16]. Note that $fb_y(\delta)$ is yet to be defined as it depends on the learning rule, i.e. on how the weight update will alter the network. Two alternatives are suggested in Equations (6) and (8).

In the attentional phase, the attentional terms β are optimized to fit target output ϵ_s for a fixed number of iterations $t = 0, \ldots, T$, called the *attention span*, after which a reward r is received. The networks reward prediction error $r - \sigma_s$ determines whether learning will enforce or discourage attention-modulated network activity. Specifically, if s is correct reward $r = 1$ will alter in the weight update in the direction of the attentional gain, while $r = 0$ in incorrect trials will update weights in the opposite direction.

Weight update. We offer two different theories on attention-guided learning. The first states that synaptic strength is updated according to the postsynaptic attentional gain while the second claims that learning is dependent on the presynaptic activity as well as relative postsynaptic attentional gain.

1. Attentional weight gain. Suppose we write the attentional term as a weight update over the inputs, i.e.

$$\begin{aligned}(1+\beta_y)\cdot y_{in} &= (1+\beta_y)\cdot \sum\nolimits_{x\in Y^{n-1}} w_{x,y}\, x_{out} + b_y \\ &= \sum\nolimits_{x\in Y^{n-1}} (1+\beta_y)\, w_{x,y}\, x_{out} + (1+\beta_y)\, b_y \\ &= \sum\nolimits_{x\in Y^{n-1}} (\Delta w^{ATT}_{x,y} + w_{x,y})\, x_{out} + (\Delta b^{ATT}_{y} + b_y),\end{aligned}$$

where $\Delta w^{ATT}_{x,y} = \beta_y\, w_{x,y}$, and $\Delta b^{ATT}_{y} = \beta_y\, b_y$. Then these weight and bias updates preserve the neural output of y_{out} following the attentional phase. When scaling the updates linear to the prediction error they become

$$\Delta w^{ATT}_{x,y} = (r-\sigma_s)\,\beta_y\, w_{x,y}, \qquad \text{and} \qquad \Delta b^{ATT}_{y} = (r-\sigma_s)\,\beta_y\, b_y. \tag{5}$$

Note that the maximum attentional gain of 50% forces $|\beta_y| \leq 0.5$. Seeing as $|r-\sigma_s| < 1$, the weight update does not exceed the weight value, i.e.

$$|\Delta w^{ATT}_{x,y}| \leq 0.5 \cdot |w_{x,y}| \leq |w_{x,y}|.$$

Therefore, weights cannot change sign.

These updates ensure that the activation of neuron y either moves into the direction of its attentional gain in rewarded trials, and in the opposite direction in unrewarded trials, i.e.

$$\begin{aligned}y_{out} &= g\Big(\sum\nolimits_{x\in Y^{n-1}} (\Delta w^{ATT}_{x,y} + w_{x,y})\, x_{out} + (\Delta b^{ATT}_{y} + b_y)\Big) \\ &= g\Big(\sum\nolimits_{x\in Y^{n-1}} ((r-\sigma_s)\,\beta_y\, w_{x,y} + w_{x,y})\, x_{out} + ((r-\sigma_s)\,\beta_y\, b_y + b_y)\Big) \\ &= g\Big(\sum\nolimits_{x\in Y^{n-1}} \lambda_{r,\beta_y}\, w_{x,y}\, x_{out} + \lambda_{r,\beta_y}\, b_y\Big) \\ &= g(\lambda_{r,\beta_y}\cdot y_{in}),\end{aligned}$$

where $\lambda_{r,\beta_y} = (r-\sigma_s)\,\beta_y + 1$. For instance, if $\sigma_s = 0.5$, the activation of a ReLU-activated neuron y with $\beta_y = 0.5$ becomes factor $\lambda_{1,0.5} = 1.25$ times its original pre-attentive value in a rewarded trial and $\lambda_{0,0.5} = 0.75$ times its original value in an unrewarded trial. For $\beta_y = -0.5$, the activation of y becomes $\lambda_{1,-0.5} = 0.75$ in rewarded trials and $\lambda_{0,-0.5} = 1.25$ in unrewarded trials.

As the weight update alters the attention-modulated network activity with a set rate, the attentional terms can be optimized with the same feedback algorithm used by supervised error-backpropagation. Therefore, the feedback $fb_y(\delta)$ on error δ received in neuron $y \in Y^n$ for $n < N$ equals

$$fb_y(\delta) = g'(y) \sum_{z\in Y^{n+1}} w'_{z,y}\, fb_z(\delta) \tag{6}$$

and $fb_y(\delta) = \delta_y\, g'(y)$ if $n = N$. Here, $g'(y)$ equals the activation function derivative, which equals the indicator function $\mathbb{1}_{\{y_{out}>0\}}$ for a ReLU-activated neuron y.

2. Attention-modulated Hebbian plasticity. Recall that the plasticity rule from Equation (2) states that learning depends on the presynaptic activity as well as the feedback received in the postsynaptic neuron concerning the prediction error. This is not the case for the attentional weight gain rules from Equation (5). Here we derive learning rules that account for this presynaptic activity.

The weight update from Equation (2) and bias update $\Delta b_y = \alpha f b_y(\delta)$ can be derived from the attention update by $\Delta w_{x,y} = \Delta\beta_y \frac{x_{out}}{y_{in}}$ and $\Delta b_y = \Delta\beta_y \frac{1}{y_{in}}$. As the attentional terms are optimized over a number of iterations during the attentional phase, this series of attentional updates $\Delta\beta_y$ can be approximated by the total attentional gain β_y. Moreover, when scaling the resulting updates linear to the prediction error, they become

$$\Delta w_{x,y}^{ATT} = (r - \sigma_s)\beta_y \frac{x_{out}}{y_{in}}, \qquad \text{and} \qquad \Delta b_y^{ATT} = (r - \sigma_s)\beta_y \frac{1}{y_{in}}. \tag{7}$$

These rules state that synaptic weights change according to presynaptic activity as well as the postsynaptic attentional gain relative to its cumulative input. Hence, these rules can be considered as attention-modulated Hebbian plasticity.

In contrast to the attentional weight gain rule, the attention-modulated Hebbian plasticity rules allow weights to change sign. As a result, attention-guided learning differs from optimizing the attentional terms using the feedback mechanism from Equation (6). For instance, for two neurons x, y with a positive connection $w_{x,y}$, the feedback mechanism in Equation (6) increases attentional activity in both neurons if y receives positive feedback. However, whether the weight update changes the sign of $w_{x,y}$ in an unrewarded trial, is not accounted for in weight updates of connections to x, resulting in sub optimal or undesirable effects on y. As a result, a feedback mechanism should account for these weight changes when optimizing the attentional gain.

The attention-modulated learning rule alters the attentional activity with rate $r - \sigma_s$ when it is optimized using the following feedback mechanism in the hidden layer

$$fb_y(\delta) = g'(y) \sum_{z \in Y^{n+1}} \left(w_{y,z} + \beta_z \frac{y_{out}}{z_{in}} \right) fb_z(\delta). \tag{8}$$

Hence, optimizing attention using this feedback rule will optimize the attentional weight update. Note that it is assumed here that the feedback and feedforward weights share the same values, i.e. $w_{y,z} = w'_{z,y}$. However, this feedback rule can be easily extended to a system where this is not the case by replacing $w_{y,z} + \beta_z \frac{y_{out}}{z_{in}}$ with $w'_{z,y} + \beta_y \frac{z_{out}}{y_{in}}$.

Approximation of Attention-modulated Hebbian plasticity. As is detailed in section 5, an approximation of the attention-modulated Hebbian plasticity rule

that minimizes the use of computationally expensive operations may increase classification performance. As operations such as multiplication accumulate rounding errors in the attentional term β_y of neuron y, its approximation $\hat{\beta}_y$ may be defined by the update

$$\Delta\hat{\beta}_y = \alpha\, fb_y(\delta).$$

Note that the update of β_y from Equation (4) can be derived from this update by $\Delta\beta_y = \Delta\hat{\beta}_y \cdot y_{in}$. Hence, by substituting $\hat{\beta}_y\, y_{in}$ for β_y in the forward pass, we derive neural output $\widehat{y_{out}}$ that approximates y_{out} by

$$\begin{aligned} \widehat{y_{out}} &= g\left((1 + \hat{\beta}_y y_{in}) \cdot y_{in}\right) \\ &\approx g\left((1 + \beta_y) \cdot y_{in}\right) = y_{out} \end{aligned} \tag{9}$$

Similarly, substituting $\hat{\beta}_z\, z_{in}$ for β_z in the feedback mechanism derives feedback $\widehat{fb_y(\delta)}$ that approximates $fb_y(\delta)$ from Equation (8) by

$$\begin{aligned} \widehat{fb_y(\delta)} &= g'(y) \sum_{z \in Y^{n+1}} \left(w_{y,z} + \hat{\beta}_z\, y_{out}\right)\, \widehat{fb_z(\delta)} \\ &= g'(y) \sum_{z \in Y^{n+1}} \left(w_{y,z} + \hat{\beta}_z z_{in} \frac{y_{out}}{z_{in}}\right)\, \widehat{fb_z(\delta)} \\ &\approx g'(y) \sum_{z \in Y^{n+1}} \left(w_{y,z} + \beta_z \frac{y_{out}}{z_{in}}\right)\, fb_z(\delta) = fb_y(\delta) \end{aligned} \tag{10}$$

Note that this approximated feedback mechanism further minimizes the use of computationally expensive division.

4 Experiments

In the previous section, two attention-guided learning rules were derived using different assumptions.

1. Attentional weight gain in Equation (5) states that synaptic strength is updated according to the postsynaptic attentional gain
2. Attention-modulated Hebbian plasicity in Equation (7) states that update is dependent on the presynaptic activity as well as relative postsynaptic attentional gain.

The classification performance of these learning rules are tested on the MNIST dataset, which consists of $70,000$ 28x28 pixel grayscale images of handwritten digits. Moreover, successful learning rules are tested on CIFAR10 and CIFAR100, which consist of $60,000$ 32x32x3 pixel colour images of objects and organisms. CIFAR10 contains 10 classes of $6,000$ images per class, while CIFAR100 contains 100 classes of 600 images per class. All sets contain a preselected test set of $10,000$ samples. Computationally expensive tests are performed on the CIFAR10 vehicle

subset, which we call CIFAR4 and consists of classes 'automobile', 'airplane', 'ship' and 'truck'. CIFAR4 therefore contains $20,000$ training samples and $4,000$ test samples. For all tests $1,000$ random training samples are selected as a validation set to measure validation accuracy during training.

How attention aids learning is reflected by the classification performance for different attention spans T. To accurately compare these for the same global rate $\hat{\alpha}$, we set the attention rates α to $\hat{\alpha}/T$. Both learning rules are tested on MNIST for several attention spans using a small neural network consisting of 3 fully connected ReLU-activated hidden layers and a linear output layer. Each of the hidden layers decrease in size from 1500 to 1000 to 500 nodes to aggregate feedforward information. Initial weights are sampled from a normal distribution with $\mu = 0$ and $\sigma = 0.05$. The goal for this experiment is to examine how attention-guided learning rules perform on simple tasks rather than to simulate biological processes.

Visual learning can be simulated in convolutional networks since these retain spatial information similar to the visual cortex. Therefore, samples of CIFAR4, CIFAR10 and CIFAR100 are fed into two consecutive convolutional layers with 32 3x3 filters. Here, the second convolutional layer aggregates information from the previous layer with stride 2 using a zero-padding. Neural output feeds into a fully connected layer of 500 nodes, followed by a linear output layer. All hidden layers are ReLU-activated and a dropout rate of 0.8 and 0.3 is applied to the second and third layer to prevent overfitting. Initial weights are sampled from a normal distribution with $\mu = 0$ and $\sigma = 0.02$.

Lastly, to verify whether attention biases the network towards relevant information while ignoring irrelevant information, we visualize network heatmaps using our implementation of gradient-weighted class activation mapping or GRAD-CAM [30]. GRAD-CAM highlights regions that contribute to a prediction by visualizing its gradient with respect to the last convolutional layer, which is believed to contain high-level spatial information. In contrast to other visualization methods such as saliency maps, GRAD-CAM is class-discriminative and is therefore more informative on model behavior.

5 Results

All results reported in this section represent the learning rule performance over five different network initializations. Each network was trained for a fixed number of epochs that increase with task difficulty, i.e. 100 on MNIST, 150 on CIFAR10 and 230 on CIFAR100. The validation set was used to ensure the model does not overfit the training data, i.e. does not contain declining sequences longer than 25 iterations.

Weight gain vs. Attention-modulated Hebbian plasticity. As listed in Table 7, the weight gain rule (WG) was considerably outperformed on MNIST by both supervised error-backpropagation (EBP) and the attention-modulated Hebbian rule (HEB) for attention spans $T = 1$, 2 and 5 with $\alpha = 0.01/T$. In contrast, HEB performed remarkably close if not identical to EBP on classifying new data for attention spans $T = 1$ and 2. Here, HEB requires $1.2 - 1.5$ times more epochs to reach maximum performance than EBP. Similarly, as can be seen in Figure 8, the learning process of HEB for these attention spans showed strong similarities to EBP while WG is outperformed on all performance measures. The performance of HEB quickly deteriorated for a higher attention span of $T = 5$. This behavior is further examined on CIFAR10.

		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	$\mathbf{98.67 \pm 0.04}$	29 ± 23

Figure 7: Test accuracy of WG and HEB on new samples for different attention spans T on MNIST. Bold text represents the best performance.

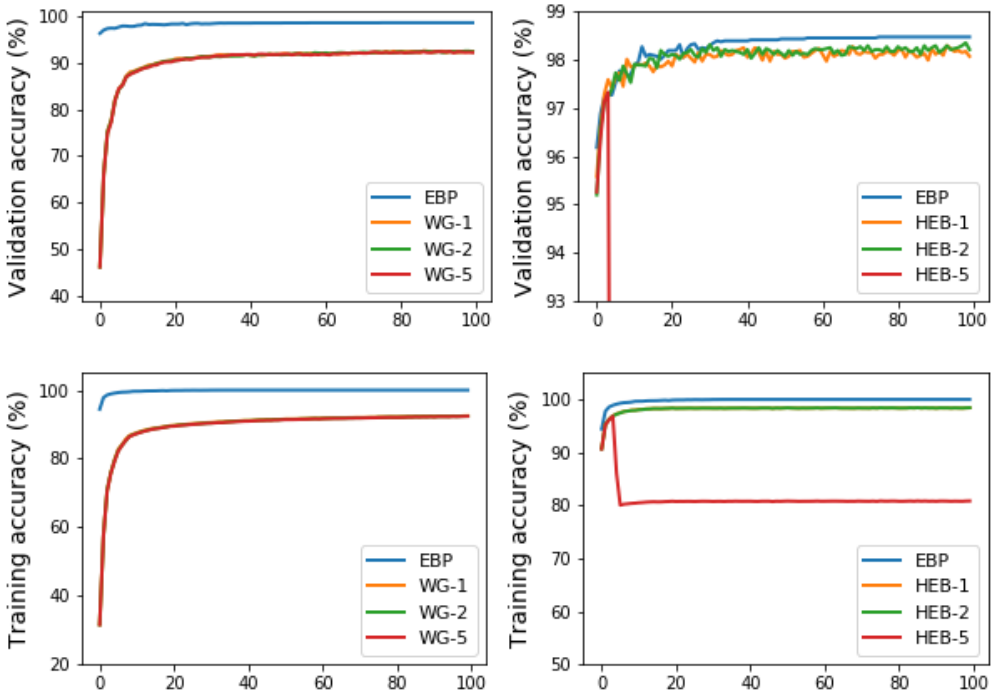

Figure 8: Validation accuracy (*upper row*) and training accuracy (*lower row*) of WG (*left*) and HEB (*right*) for different attention spans T on MNIST, signified by WG-T and HEB-T.

Attentional performance gains. Since HEB performed comparably to EBP for low attention spans on simple tasks, HEB is used to simulate attention-guided learning on CIFAR10 and its vehicle subset. As listed in Figure 9, HEB was slightly outperformed by EBP when classifying new CIFAR10 samples for attention spans $T = 1$, 2 and 5 with $\alpha = 0.005/T$. However, as illustrated in Figure 10, learning became rapidly unstable with T on the smaller set CIFAR4. This instability can be traced back to an accumulation of rounding errors in early layers. Here, the division and multiplication of small numbers, e.g. in the order of 10^{-12}, causes rounding errors, specifically when computing the relative attentional gain in the feedback mechanism[1]. The accumulation of these errors during the attentional phase disrupts weight updates, causing a steep decline in performance.

As detailed in section 3, an approximation rule of HEB, which we refer to as APPROX, minimizes the use of these operations. As detailed in Figure 9, APPROX performed similar if not better than HEB in classifying new CIFAR10 samples while even outperforming EBP for $T = 5$. Moreover, learning did not deteriorate on CIFAR4, although APPROX was slightly outperformed by EBP when classifying new samples for attention spans T=$1, 2, 5, 10$ and 20.

Performance measures from Figure 9, Figure 10 and Figure 17 from Appendix 7.1 indicate that while attention possibly increases classification performance, our results do not prove significant attentional performance gains. This is further confirmed by examining the attentional gain and neural output measures in the attentional phase of a random sample during training on CIFAR10, see Figure 13 and 18. These results indicate that attentional gain, and therefore neural output, is a linear function of the number of iterations in the attention phase. As a result, attention-guided learning for attention span T and rate α/T simulates learning with span 1 and rate α. In section 6 we will argue that this is likely a result of network parameters, specifically a low attention rate with respect to the size of the weight update.

Previous results indicate no significant attentional performance gain of APPROX on CIFAR10 and CIFAR4. As a result, testing of APPROX on CIFAR100 was limited to attention span $T = 1$ using $\alpha = 0.005$. Moreover, the performance of APPROX-1 is compared to Q-AGREL [23] with batch size 1 and $\alpha = 0.01$. As EBP overfitted the training data for long training phases, training was stopped at epoch 180. As can be seen from Figures 11 and 12, EBP outperforms both APPROX-1 and Q-AGREL due to the large amount of available feedback information, i.e. 100 times more than in trial-and-error learning. Moreover, APPROX-1 outperforms Q-AGREL in classifying both training samples and new samples. This is most evident for training samples, as APPROX-1 backpropagates loss over all output nodes, and is therefore able to learn at a faster rate than Q-AGREL, which backpropagates loss over a single output node.

[1] i.e. for feedback to neuron y; $\beta_y \frac{x_{out}}{y_{in}}$

		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	$\mathbf{72.34 \pm 0.18}$	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	$\mathbf{86.21 \pm 0.51}$	134 ± 9

Figure 9: Test accuracy of HEB and its approximation APPROX for different attention spans T on CIFAR10 and its vehicle subset CIFAR4. Bold text represents the best performance on a dataset.

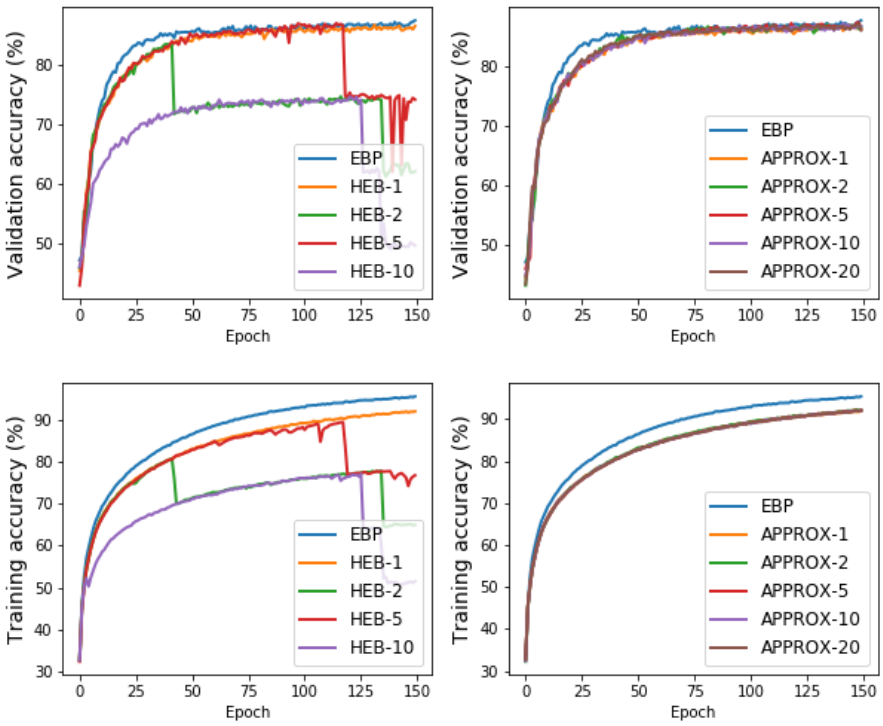

Figure 10: Validation (*upper row*) and training accuracy (*lower row*) of HEB (*left*) and its approximation APPROX (*right*) on CIFAR4 for different attention spans T, signified by HEB-T and APPROX-T.

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

Figure 11: Test accuracy of APPROX-1, EBP and Q-AGREL on CIFAR100.

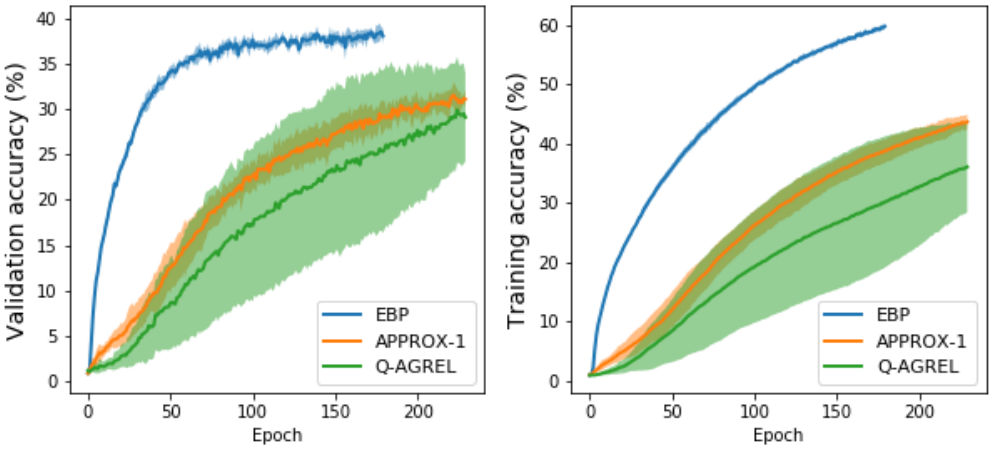

Figure 12: Validation (*left*) and training accuracy (*right*) of APPROX-1, EBP and Q-AGREL on CIFAR100.

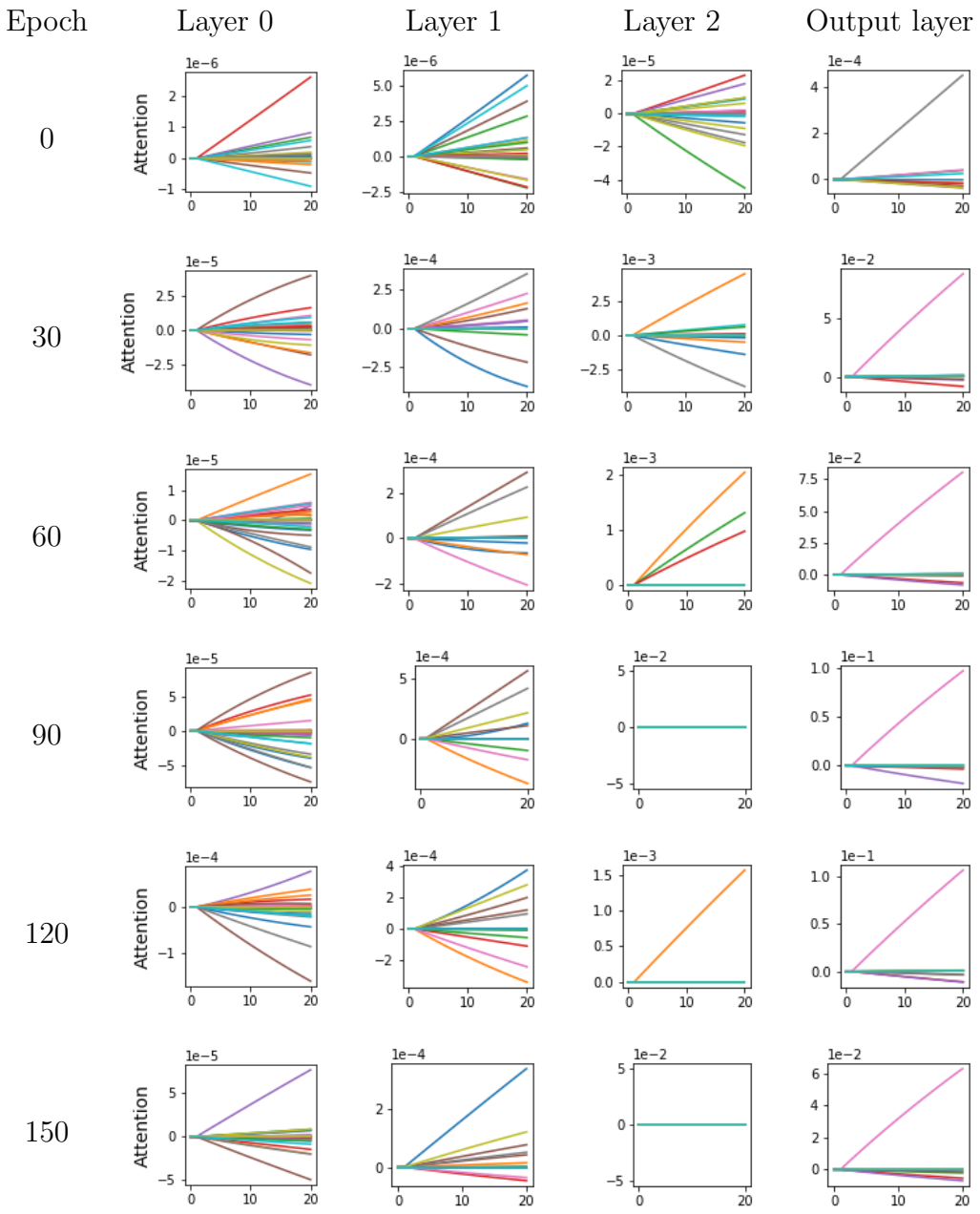

Figure 13: Attentional terms β of 40 neurons from each layer during the attentional phase when training on CIFAR10 with $\alpha = 0.0005$.

Visualizing attentional bias As was explained in section 4, GRAD-CAM highlights regions that contribute to a prediction by backpropagating its gradient to the last convolutional layer [30]. Here, gradients are averaged over the width and height dimensions to make up an importance weighting a_k^c of each activation map k for prediction c. The a_k^c-weighted sum of neural output $A(x, y, k)$ over all activation maps k signifies the importance of each (x, y) location in the final convolutional layer. The ReLU function is applied to the resulting heatmap to only visualize areas that have a positive effect on c, i.e. the importance of (x, y) to prediction c equals

$$\text{GRAD-CAM}^c(x, y) = ReLU\left(\sum_k a_k^c \, A(x, y, k)\right)$$

Figure 14, 15 and 16 show heatmaps that highlight important regions for a prediction during training using GRAD-CAM. No shift towards relevant regions indicating attentional bias can be seen during the attentional phase. This finding coincides with the notion that the attentional rate is likely too low to cause attentional performance gain. However, a clear shift towards relevant regions can be observed during training. For instance, the background in the sample from Figure 14 provokes the network to predict the 'bird' class during the first epoch. In a later training stage, the network has learned to predict the correct label by attending the horses' hind legs and equestrian. Similarly, the network predicts the automobile from Figure 15 by attending the car door and frame, and learns that the airplane from Figure 16 can be recognized by its wing and shark fin tail.

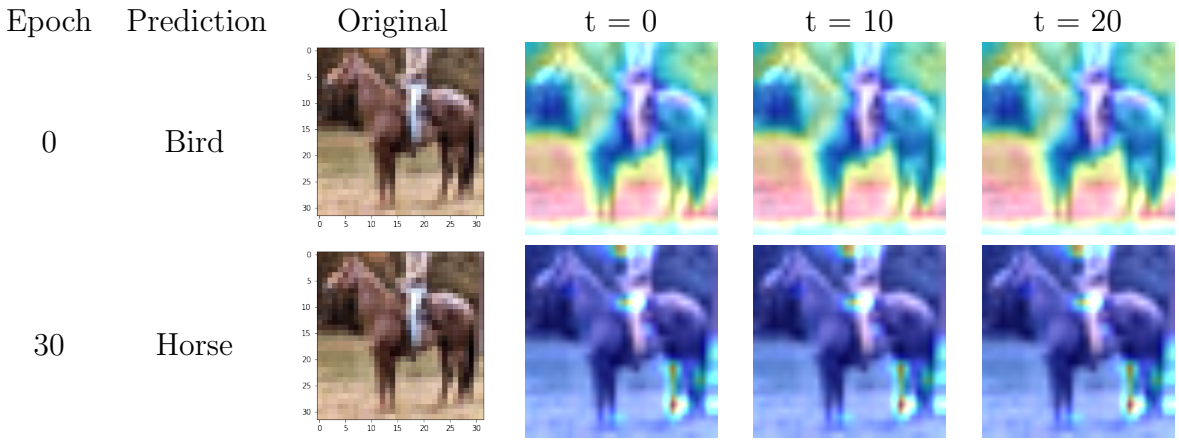

Figure 14: Heatmaps of a training sample of a horse during the attentional phase

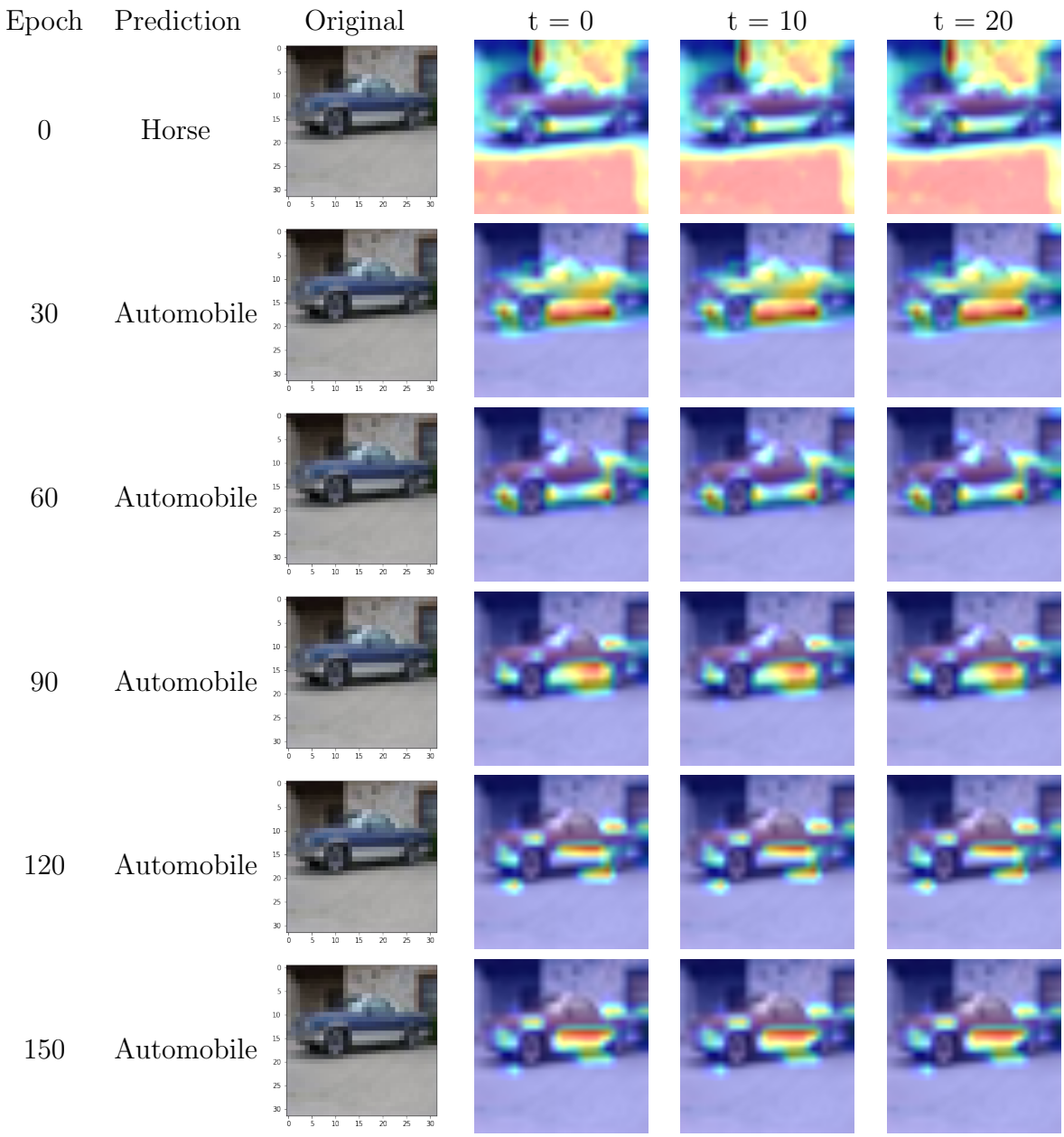

Figure 15: Heatmaps of a training sample of an automobile during the attentional phase

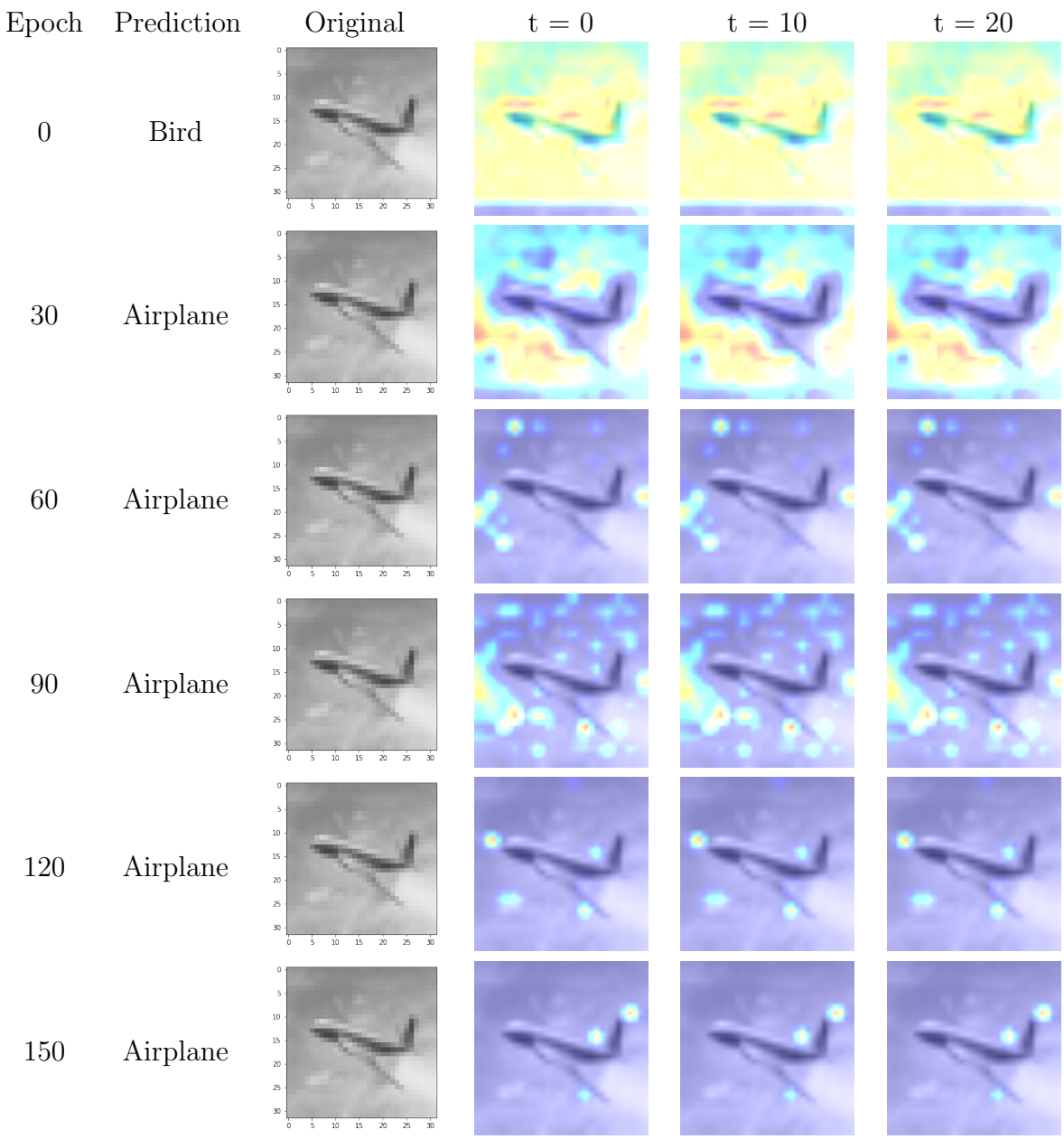

Figure 16: Heatmaps of a training sample of an airplane during the attentional phase

6 Discussion

Our results show that the HEB learning rule that alters synaptic strength according to presynaptic activity and the relative postsynaptic attentional gain, performs similarly to supervised error-backpropagation for short attention spans. Furthermore, a computationally cheaper approximation APPROX of this learning rule extends this finding to larger attention spans and even outperforms supervised error-backpropagation in some cases. This finding is most evident on sets that contain a large number of training samples per class such as CIFAR10. Moreover, our results show that the learning rule WG that scales synaptic strength according to the postsynaptic attentional gain is considerably outperformed by error-backpropagation, even on a simple task. However, as restrictions on the size of the attentional gain prevent weights from changing their sign, this learning rule may perform better using a larger network architecture.

While attention generally increased the performance of APPROX, our results did not indicate significant attentional performance gains. An examination of the network activity during the attentional phase suggests that attention is generally updated with a small, set rate. That is, attentional gain may be a linear function of the attention span T and attention-guided learning therefore equals learning with span 1 when using the same global attention rate α. As this rate α resembles the size of the weight update, it is very small. As a result, attentional updates cause no significant changes in the neural output or consequently, the loss derivative that would lead to attentional performance gains. However, scaling the size of the weight update with a learning rate $\alpha_w < \alpha$, e.g. $\alpha_w = 0.1\alpha$, would allow larger attention rates that could result in attentional performance gains. This addition to the APPROX learning rule could result in attentional performance gains that may outperform error-backpropagation.

While this study of attention-guided learning rules has been performance oriented so far, our findings coincide with well-established neurophysiological beliefs on synaptic plasticity. Hebbian theory states that synaptic strength is altered according to the causality between presynaptic and postsynaptic firing [26], i.e. "cells that fire together, wire together". Neural firing modulations, such as the attentional response gain observed in the visual cortex, can therefore be captured in a synaptic weight update according to Hebbian plasticity. By describing Hebbian plasticity as a function of the postsynaptic attentional gain, this study is the first to explain neural modulations observed in the brain before a reward is received. However, our study is limited to the case where feedback and feedforward connections share the same weights. Contrarily, prediction errors in the brain are backpropagated through a dopaminergic system [27] with separate feedback weights. Therefore an extension to the attention-modulated Hebbian learning rule that simulates the biological feedback mechanism more closely is described in section 3.

In this study, we argued that current attentional models either do not reflect biological learning mechanisms in the visual cortex or fail to explain how learning relates

to known attentional modulations. The learning framework proposed in this study is the first to describe and simulate Hebbian plasticity as a function of attentional response gain. Moreover, our results suggest that these attention-guided reinforcement learning rules result in remarkably efficient classification. This study proposes that the attention-modulated Hebbian learning framework reflects the mechanism behind attention-guided learning, which ultimately governs behavior.

References

[1] Jimmy Ba, Volodymyr Mnih, and Koray Kavukcuoglu. Multiple object recognition with visual attention. *CoRR*, abs/1412.7755, 2014.

[2] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate, 2014. cite arxiv:1409.0473Comment: Accepted at ICLR 2015 as oral presentation.

[3] Daniel Baldauf and Robert Desimone. Neural mechanisms of object-based attention. *Science*, 344(6182):424–427, 2014.

[4] Farhan Baluch and Laurent Itti. Mechanisms of top-down attention. *Trends in Neurosciences*, 34(4):210 – 224, 2011.

[5] Irwan Bello, Barret Zoph, Ashish Vaswani, Jonathon Shlens, and Quoc V. Le. Attention augmented convolutional networks. *CoRR*, abs/1904.09925, 2019.

[6] Marisa Carrasco. Visual attention: The past 25 years. *Vision Research*, 51:1484–1525, 2011.

[7] Maurizio Corbetta and Gordon L. Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3:201–215, 2002.

[8] Robert Desimone and John Duncan. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(1):193–222, 1995. PMID: 7605061.

[9] Barbara Anne Dosher and Zhong-Lin Lu. Noise exclusion in spatial attention. *Psychological Science*, 11(2):139–146, 2000. PMID: 11273421.

[10] Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, and Yann N. Dauphin. Convolutional sequence to sequence learning. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, pages 1243–1252. JMLR.org, 2017.

[11] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross B. Girshick. Mask R-CNN. *CoRR*, abs/1703.06870, 2017.

[12] Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman, and Phil Blunsom. Teaching machines to read and comprehend. In *Advances in Neural Information Processing Systems 28*, pages 1693–1701. 2015.

[13] Grace W. Lindsay and Kenneth D. Miller. How biological attention mechanisms improve task performance in a large-scale visual system model. *bioRxiv*, 2018.

[14] Sam Ling and Marisa Carrasco. Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research*, 46(8):1210 – 1220, 2006.

[15] Julio Martinez-Trujillo and Stefan Treue. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current biology : CB*, 14:744–51, 06 2004.

[16] John Maunsell. Neuronal mechanisms of visual attention. *Annual Review of Vision Science*, 1:373–391, 11 2015.

[17] John H.R. Maunsell and Stefan Treue. Feature-based attention in visual cortex. *Trends in Neurosciences*, 29(6):317 – 322, 2006. Neural substrates of cognition.

[18] Volodymyr Mnih, Nicolas Heess, Alex Graves, and Koray Kavukcuoglu. Recurrent models of visual attention. In *Advances in Neural Information Processing Systems 27*, pages 2204–2212. 2014.

[19] J Moran and R Desimone. Selective attention gates visual processing in the extrastriate cortex. *Science*, 229(4715):782–784, 1985.

[20] Amy M. Ni and John H. R. Maunsell. Spatially tuned normalization explains attention modulation variance within neurons. *Journal of Neurophysiology*, 118(3):1903–1913, 2017.

[21] Amy M. Ni and John H.R. Maunsell. Neuronal effects of spatial and feature attention differ due to normalization. *Journal of Neuroscience*, 39(28):5493–5505, 2019.

[22] Marius Peelen and Sabine Kastner. Attention in the real world: Toward understanding its neural basis. *Trends in cognitive sciences*, 18, 03 2014.

[23] Isabella Pozzi, Sander M. Bohte, and Pieter R. Roelfsema. A biologically plausible learning rule for deep learning in the brain. *CoRR*, abs/1811.01768, 2018.

[24] John H. Reynolds and David J. Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 1 2009.

[25] Pieter Roelfsema and Anthony Holtmaat. Control of synaptic plasticity in deep cortical networks. *Nature Reviews Neuroscience*, 19:166–180, 02 2018.

[26] P.R. Roelfsema and A. van Ooyen. Attention-gated reinforcement learning of internal representations for classification. *Neural Computation*, 17(10):2176–2214, 2005.

[27] J.O. Rombouts. *Biologically Plausible Reinforcement Learning*. PhD thesis, Vrije Universiteit Amsterdam, 2015. Aard- en Levenswetenschappen.

[28] Douglas Ruff and Marlene Cohen. A normalization model suggests that attention changes the weighting of inputs between visual areas. *Proceedings of the National Academy of Sciences*, 114:201619857, 05 2017.

[29] Katharina N. Seidl, Marius V. Peelen, and Sabine Kastner. Neural evidence for distracter suppression during visual search in real-world scenes. *Journal of Neuroscience*, 32(34):11812–11819, 2012.

[30] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, 128(2):336–359, Oct 2019.

[31] Sainbayar Sukhbaatar, Edouard Grave, Guillaume Lample, Hervé Jégou, and Armand Joulin. Augmenting self-attention with persistent memory. *CoRR*, abs/1907.01470, 2019.

[32] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems 30*, pages 5998–6008. 2017.

7 Appendix

7.1 Performance measures

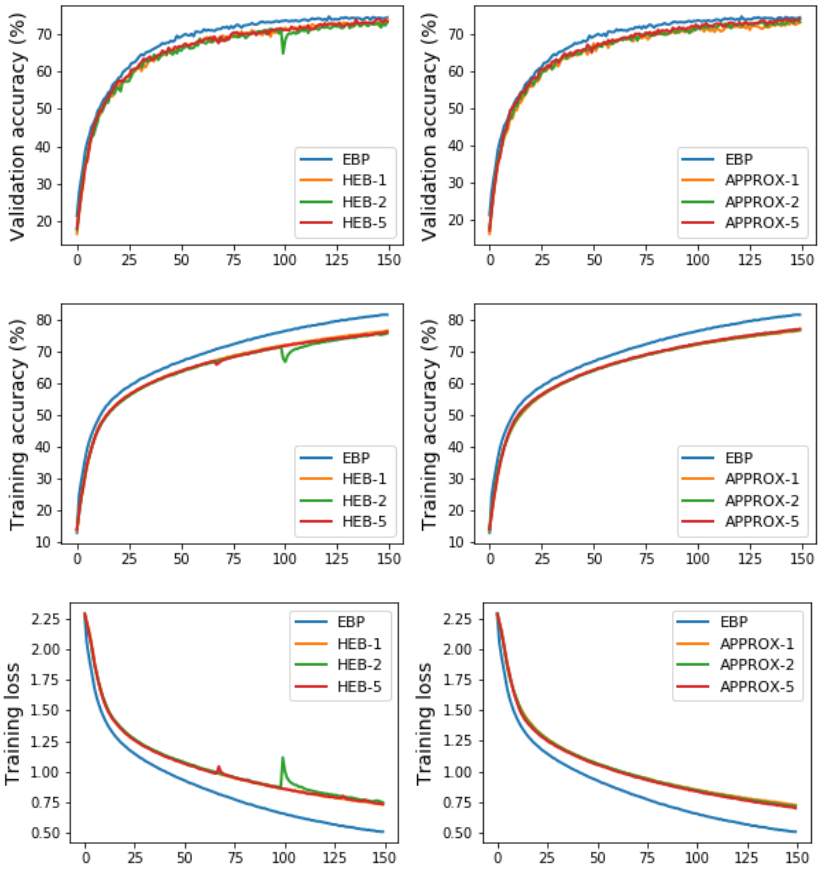

Figure 17: Validation accuracy (*upper row*), training accuracy (*middle row*) and loss (*lower row*) of HEB (*left*) and APPROX (*right*) on CIFAR10 for different attention spans T, signified by HEB-T and APPROX-T.

7.2 Attentional gain

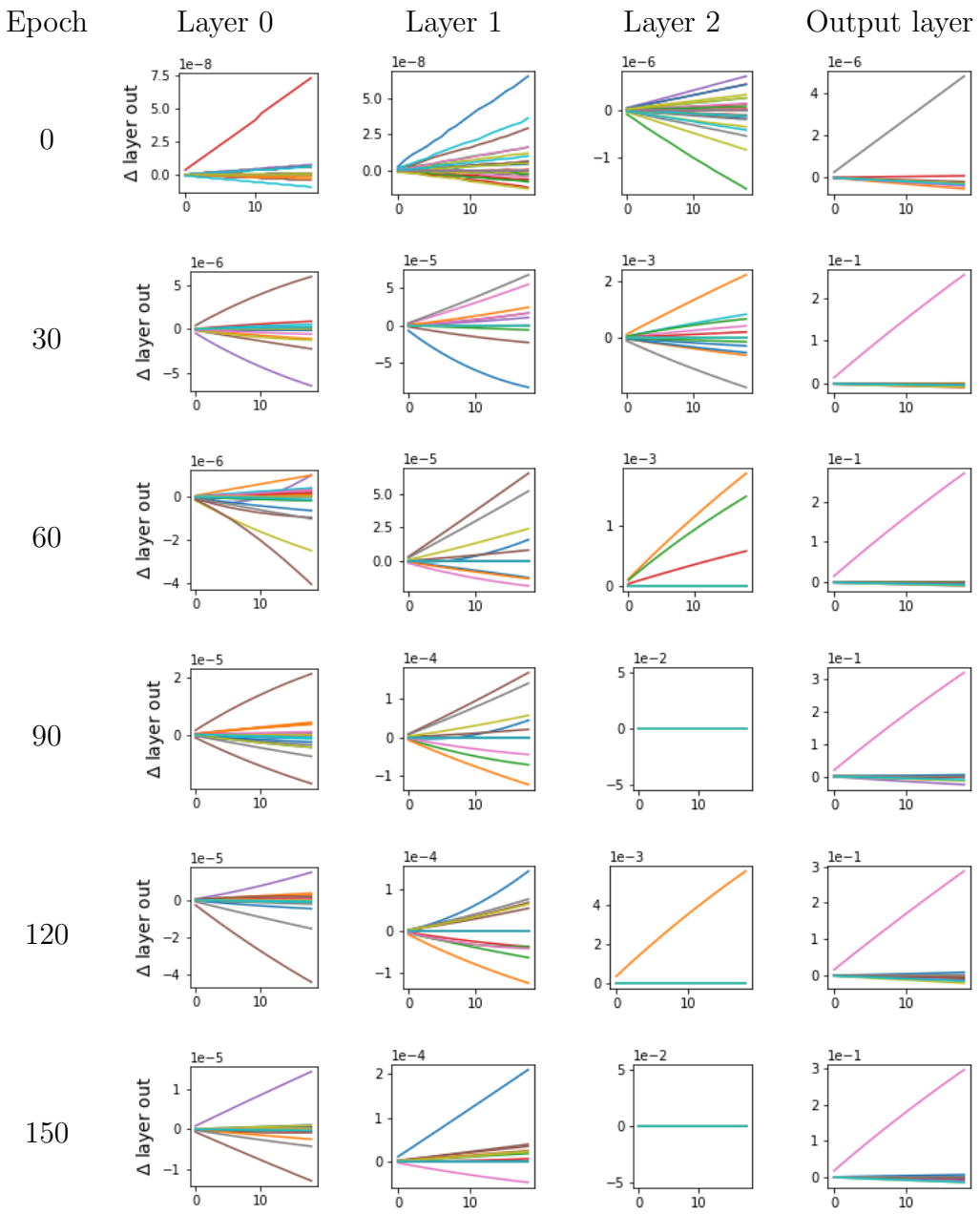

Figure 18: Attentional response gain of 40 neurons from each layer during the attentional phase when training on CIFAR10 with $\alpha = 0.0005$.