Image and pattern reconstruction using differentiable plasticity

Jigalur Priyanka¹, B G Prasad²

¹Department of Computer Science & Engineering, B.M.S. College of Engineering, Bengaluru, India
²Department of Computer Science & Engineering, B.M.S. College of Engineering, Bengaluru, India
Email: priyanka.jigalur@gmail.com
bgprasad.cse@bmsce.ac.in

Abstract. The brain is a substantial boon to humankind that adapts nature accordingly. The brain can learn and unlearn based on the situation. This singularity of human learning led to the research creating models using Artificial Intelligence (AI) to incorporate the brain’s behavior. The investigation opened up many new approaches to study AI with neural networks by adding new techniques to imitate the human brain’s functionalities. Many models can learn from experience like Recurrent Neural Network (RNN’s), Long Short Term Memory (LSTM) with the fixed network size. This paper describes the simple method of creating the model which will behave similar to the biological brain and recreates its differentiable plasticity to adopt the features of neural network connection. It also shows that applying plasticity and the Hebbian plastic connection rule can result in optimization in RNN. This new approach of reconstruction of images based on plastic neural network experiments shows that the above novel approach gives more optimized results than the traditionally used RNN techniques. In this paper, a proposal is made where models can memorize and reconstruct unseen sets of images by solving recurrent networks using plasticity rules.

Keywords— Differentiable Plasticity; Artificial Intelligence; Long Short Term Memory; Image Reconstruction; plasticity; Neural Networks; Hebbian Trace; Meta-Learning.

1. Introduction

Advancements in the AI domain have paved the way for improving existing applications, platforms that rely on cutting-edge AI. A similar area of research where machine learning and neuroscience play a significant role in building different brain-like networks.

Machine learning is a technology industry in AI that allows the systems to learn from the past and develop the model without programming explicitly [1]. Machine learning involves developing a model that can independently access and learn from past data. Over the years, numerous meta-learning algorithms have suggested that models mainly try to learn from their experiences [2]. On the other hand, these models were never more effective at adapting and functioning like the human brain, applying what it has learned in the past to current situations. Examining how the brain learns through synaptic plasticity will help to illustrate previous learning experiences. The activity simultaneously activated and linked
with neurons during tasks strengthens the network which is created during that task. Hebb rule is defined as neurons which fire together then they wire together when the synaptic plastic rule is imposed.

Mathematical models designed previously were all illustrative illustrations of images of humans, called handcrafted models [3]. These were later replaced by developing data-driven models, and these models are used till now. On the other hand, observed data make sure that they are part of the learning of the model. Deep learning models are now molded on data-driven learning instead of hypothetical human designs when resources and vast data are computed. Various deep learning methods are developed to extract and adapt relevant information from enormous data sets. This learning paradigm is increasingly being used to study the neural network complexity in meta-learning. Meta-learning is the most challenging application and the most significant problem in machine learning because it provides insight into all past, current, and future tasks. Although training a single neural network is complex, the computations based on the knowledge are designed to mimic the brain of a human.

Furthermore, the architecture of supervised meta-learning might be applied to several shots, a fascinating classification phenomenon. This architecture solves deep learning challenges, and learning the agent's knowledge has made significant progress. Since every advancement has limitations, present models require comprehensive retraining and many training samples to succeed at various tasks using traditional approaches. Additionally, these models are learned using massive datasets that fail to update weights without forgetting previously obtained knowledge using backpropagation approaches, resulting in a lengthy learning process. This is a deterministic feature of the network that leads to the slow learning of the model. The research contains information taken from the original images that can recreate the image features as ground truth to integrate into a superior neural network all the limitations of profound learning models.

With the help of backpropagation, a new learning paradigm was devised to emulate biological neurons. It uses cellular neuromodulation as inspiration to create a new deep neural network model, which is explicitly designed in order to learn adaptive behaviours. Plasticity is defined as a biological term that describes the human brain’s ability to change over time by creating new connections between neurons and destroying those that are no longer needed. However, synaptic weights are automatically modified in response to ongoing activity. The development of artificial neural networks in neural plasticity, or the weakening and strengthening of connections between neurons based on brain activity, have been widely studied. The fundamental benefit of utilizing differentiable plasticity is that it allows for gradient descent training of plastic link behaviour so that previously trained networks can respond to future conditions. Thus, neural plasticity is defined as the evolution of memory access mechanisms in plastic neuronal memory models to give variable attention to stored information.

This concept is used in various domains, including image inpainting, editing images, and translation of the image. In the paper, experiments are focused on image reconstruction. The paper explains the approach which can train a large network for various non-trivial tasks. To use this strategy, plasticity is applied over various classes of images that memorize the pattern and reconstruct the images. The paper finally shows that the model can memorize the images and reconstruct them whenever the plasticity concept is applied to the model.

2. Literature Survey

In the field of biological neuron replication, the application of AI appears to be novel; the study has been going on for a long time. Let us take a look at how the design of building a network similar to humans has evolved. Deep learning methods have made improvisation in many fields such as speech recognition, complex tasks on unstructured data, and problems requiring high abstract feature extractions. However,
they are held back often due to their limitations, such as the combinatorial explosion in real-world datasets, adequate only to large amounts of datasets, and the explosion of gradients [4].

In 1985, AI research was motivated to replicate human intelligence, including an approach constructed to manipulate symbolic representations. Artificial intelligence is the science for researching theories and simulating the technologies which behave like the human brain. One can stimulate the brain by creating models. Meta-learning and learning-to-learn were introduced as literary theories in 1987 [5].

The advancement in meta-learning led to the construction of task-adaptive learning and meta-learning assistants. Christophe Giraud-carrier, in his paper, explained that meta-learning could be defined as the act of utilizing the knowledge which has already been learned to facilitate understanding and enhance performance [6].

Meta-learning occurs in the form of patterns across domains in the realm of inductive transfer and learning-to-learn. Invariant transformations are a general understanding of the nature of patterns across disciplines. Image recognition becomes more accessible when a target item is invariant under rotation, translation, scale, and other factors [7].

Synaptic plasticity was initially discovered at the end of the nineteenth century when the first proof of long-term continuous distinct turnovers that contribute to synaptic intensity was found; even then, it took about eighty years for this experimental evidence to be acknowledged. Graves investigated how RNNs could be used to tackle algorithmic difficulties. The authors tested their LSTM designs to determine if the results were appropriate for various meta-learner tasks that an LSTM might handle [8].

Meta-learning represents a higher level of comprehension than only information and adjustment during the learning process. Meta-learning will evaluate an individual's mindful learning style, provide the required tools, and change it based on current activities. The performance of LSTM is superior in the tasks such as regression and classification with sparse data [9].

Elia smith, in his book, proposed the architecture called “Semantic Pointer Architecture,” where the model adopts computations, representations cognitively. The architecture was influenced by understanding cognition as a biological process [10].

The machine learning models try to learn over multiple experiences, and the performance increases with an increase in learning episodes. With the advancement of research, various learning models based on dynamic programming methods have developed in artificial intelligence. In the article, the author described the learning algorithm based on dynamic programming, by which the model can enhance its long-term performance with an increased experience [10].

The areas of machine learning and neurocomputing have revolutionized long-term memory. One reason for this recurrent network's efficacy is its ability to process the exploding problem or vanishing gradient problem, which is a difficult challenge to solve when training recurrent and deep neural networks [12]. The popularity of LSTMs grew as a result of their ability to overcome the problem of vanishing gradients. However, it turns out that they did not delete it. The issue is that data must still be moved from cell to cell to be evaluated. Furthermore, with new features such as forget gates, the cell has become rather sophisticated. As a result, an efficient neural network forming a task that provides the network with the ability to transmit a complex life cycle throughout learning continues to be a significant focus of research; hence, it has many connections that must be preserved.

As a result, various methodologies have been accepted in neuroscience, which can be considered design standards for synaptic plasticity. Aptness to change present neural network procedures in the future and adapt responses to input activity and undergo synaptic plasticity alterations. The connections formed between neurons in the brain are not fixed but somewhat plastic [9]. As a result, synaptic neuron contacts in a neural network are involved in a certain way in a neural network, raising or lowering
activity. Additionally, synaptic plasticity changes the synaptic machinery; presynaptic neurons and postsynaptic neurons can be broken down as long-term memory or short-term memory [13]. Long-term synaptic plasticity consists of temporary memory. However, the memory changes over the years by forming lifelong memories, which happen at various time intervals.

Furthermore, the Hebbian theory describes neural adaptation during synaptically connected neurons, leading to long-term synaptic changes. Thus, Hebb postulate is a core idea and a fundamental process for synaptic plasticity [14]. As a result, meta-learning is associated with Synaptics, and in addition to that, it is a study of current image restoration techniques.

Deep learning has produced breakthroughs in several other fields, such as speech recognition, but using deep learning to solve image problems has emerged. The deep neural network is complicated when it comes to training. As the layers increase, the network encounters vanishing and gradient explosion problems. A degradation problem occurs when deep networks start to converge: as network depth increases, accuracy becomes saturated and deteriorates quickly. Surprisingly, overfitting does not cause this degradation because adding more layers to a sufficiently deep model increases training error.[11] Increasing more layers does not solve that problem. However, it does result in a more significant training error or network structure overfitting. As a result, a unique neural technique for image classification and meta-learning was developed.

Thomas Miconi et al.[15] in their work, they trained a neural network with backpropagation and explained Hebb’s rule. The rule is used to specify a plasticity rule that aids if a neuron participates in making another neuron fire multiple times, the link between them is reinforced. (also known as neurons that fire together then wire together). Thus, a connection between any two neurons contains a fixed and a plastic component, with the fixed component being the connection weight and the plastic component being the Hebbian trace. The results from the Hebbian path, i.e., the product of the pre-synaptic and post-synaptic activity, averaged over time. For a given task of pattern recognition, LSTM solves the task imperfectly after about 5 lakh episodes; using plastic RNN, the task was solved within 2000 episodes, 250 times faster. Further, trained plastic networks might offer an innovative solution to the challenge of learning to learn. The main advantage of using a plasticity model is that a regular neural network can remember a single image class at a time. In contrast, the plasticity model can train a multi-class to train the model with multiple image classes and predict the output for any class. One more advantage of using plasticity is that it can resize the neural networks based on the feature it is learning. In contrast, it is not possible in other neural networks where there is fix size image before training and test; moreover, differentiable plasticity makes the plasticity for each connection trainable, resulting in higher accuracy and computations than the non-differentiable plasticity method.

3. Methodology

The section shows how to design and train the model to create a neural network using a differentiable plasticity rule combined with neuromodulation techniques and the Hebbian trace.

A. Differentiable Plasticity Using Hebbian Trace

The current work extends the differentiable plasticity approach, which uses gradient descent to maximize each network’s weights and plasticity. Each network connection is strengthened by a Hebbian plastic component that decays and grows in response to activity automatically. A plastic and a fixed component are included in every connection:

\[ x_j(t) = \sigma \left( \sum_{i \in \text{inputs to } j} (w_{ij} + a_{ij} \text{Hebb}_i(t)) x_i(t-1) \right) \]  \[16\]  \[1\]
\[ \text{Hebb}_{ij}(t+1) = \eta x_i(t-1)x_j(t) + \text{Clip}(\text{Hebb}_{ij}(t)) \]  \hspace{1cm} (2)

The neuron output at time \( t \) is \( x_i(t) \) for any neuron \( i \). Nonlinear is represented with \( \sigma \) (tanh is used for calculation). Weight \( w_{ij} \) is considered the non-plastic component weight that connects different neurons such as \( i, j \), and \( \alpha_{ij} \) is the coefficient of the plasticity scale of the network’s component. \( \text{Hebb}_{ij} \) is considered the plastic component in the case of Hebbian trace, which collects neuron \( i, j \) in both postsynaptic and presynaptic activity, as shown in equation 2. At the start of each episode, \( \text{Hebb}_{ij} \) is set to zero. In the neural network, the structural components are represented as \( \eta, w_{ij} \) and \( \alpha_{ij} \), optimized by the gradient descent from one episode to another. Thus, the expected loss over one episode is minimized.

In Eq. 2, the \text{Clip}(x) \) function is the function that confines \( \text{Hebb}_{ij} \) to the range from -1 to +1, where Hebbian learning’s inherent instability is eliminated. This function was either a normalization done by Oja's rule or a simple decay term. The difference between the \( \alpha_i \) and \( \eta \) parameters: is the plastic connections' intra-life "learning rate." Here \( \alpha_{ij} \) is defined as the scale parameter that establishes the maximum magnitude of the plastic component, and it controls how rapidly new information is integrated into the plastic component (as the range of \( \text{Hebb}_{ij} \) is restricted from -1 to +1). The plastic component tries to quickly learn and update the parameters to retain the episodic memory in each episode independent from other recent events. The hidden activation layers are associated with similar classes aggregated into a single vector class in the Hebb, resulting in a compressed episodic memory reflecting distinct episodic memory traces. This process achieves the learning of unusual classes by speeding up the binding of class labels to deep data.

**B. Neuromodulation**

The neuromodulation model can be incorporated into the plasticity in this architecture. This simple approach can be achieved by adding the parameter \( \eta \), which is dependent on either one or multiple neurons in the network. In order to decide how plastic connection should be at any given time \( t \), the parameter \( \eta \) should be kept under control because it impacts the rate at which they change. In this simple neuromodulation model, the modification to the equations is to replace Eq. 2 with the network-computed, time-varying neural signal \( M(t) \). Equation 2 is replaced by

\[ \text{Hebb}_{ij}(t+1) = M(t)x_i(t-1)x_j(t) + \text{Clip}(\text{Hebb}_{ij}(t)) \]  \hspace{1cm} (3)

**4. Results**

The differentiable plasticity concept can be evaluated in various applications. This paper is validated in memorizing the coloured images, so that network tiers reconstruct the image when distorted images are given. This unconventional method is different from the algorithms such as CNN or any existing deep learning models that depict the reconstruction of images. The technique used during training is different from standard methods. In the single course of training, the model is trained over multiple classes containing many samples. In the case of testing, images are chosen at random from the trained classes.

The dataset used to experiment is the CIFAR-10 dataset. The dataset comprises 60,000 colored images of 32x32x3, categorized into ten classes, with each class containing 6,000 images. Totally out of 60,000 images, 50,000 images and 10,000 images are classified as training and testing images. The CIFAR-10 is divided in terms of batches for training and testing. There are five batches for training and one batch used for testing, with each batch size of 10,000 images. The test batch contains the sample images, which is randomly picked from each class. Figure 1 shows the different classes which are present in the CIFAR 10 data set.
Since natural images are correlated highly, neighbouring pixels in natural images most likely to have related values, which are used to create the co-occurrence matrix. Therefore, the values in the matrix are diagonally dispersed; the average of close-up pixels must be seen as a trivial approximate result. In the case of testing, half of the pixels in the image remain unchanged, which provides the information for image reconstruction in the form of the pixel, which means they will share information with the remaining test pattern. The output is then constructed using a positive-correlation diagonal matrix, in which the plasticity component duplicates the correlation generated in the image pixels, validating the following image. This is due to the fact that neighbouring pixels act as highly correlated neurons. As a result, each pixel receives content from its immediate neighbours, allowing the entire image to be reconstructed.

The training and testing of the model are done in the form of episodes using the image dataset, i.e., the CIFAR-10 dataset. In each new episode, three images are chosen, each of which can be either the same class or of a different class based on the task. With 1,025 neurons in total, the resulting in $3 \times 1,025 \times 1025 = 3,151,875$ parameters where 3 represents RGB values. Each episode featured three images that were displayed three times (in random sequence each time) for a total of 20 timesteps, with three-time steps of zero input in between. In testing the model, one image is taken randomly, halved at any angle from three images. This step is carried out three times. The model tries to reconstruct the image from the previously learned model. The final reconstructed image is shown in the last column. The experiment shows that the model has mastered the ability to recall and reconstruct images. Figure 2 shows the sample output. The optimizer used is the Adam optimizer which is an adaptive learning algorithm. The main advantage of using this optimizer is that it computes the individual learning rate for various parameters and prevents the model from gradient descent problems.
5. Conclusion

The research in the field of image reconstruction has the advancement of various learning models based on dynamic programming methods have developed in artificial intelligence. But the main disadvantage of using the traditional techniques are the neural networks are unable to learn and stores the information. The fundamental benefit of utilizing differentiable plasticity is that it allows for gradient descent training of plastic link behaviour so that previously trained networks can respond to future conditions. The differentiable plasticity technique also reduces the time taken by the model during training by continuously updating the Hebbian value throughout the training process. This model also has the ability to update the weights on every epoch based on the network learnt which in turn helps the model to increase or decrease the network size. On the whole this new approach of reconstruction of images based on plastic neural network experiments shows that the novel approach gives more optimized results than the traditionally used RNN techniques.

References

[1] Kalali, Amir, Sarah Richerson, Emilia Ouzunova, Ryan Westphal, and Bradley Miller. "Digital Biomarkers in Clinical Drug Development." In Handbook of Behavioral Neuroscience, vol. 29, pp. 229-238. Elsevier, 2019.
[2] Vanschoren, Joaquin. "Meta-learning: A survey." arXiv preprint arXiv:1810.03548 (2018).
[3] Nanni, Loris, Stefano Ghidoni, and Sheryl Brahnam. "Handcrafted vs. non-handcrafted features for computer vision classification." Pattern Recognition 71 (2017): 158-172.
[4] Deng, Li, and Dong Yu. "Deep learning: methods and applications." Foundations and trends in signal processing 7, no. 3–4 (2014): 197-387.
[5] Goertzel, B., and P. Wang. "Adaptive algorithmic hybrids for human-level artificial intelligence." Advances in Artificial General Intelligence: Concepts, Architectures, and Algorithms (2007): 94.
[6] Giraud-Carrier, Christophe, Ricardo Vilalta, and Pavel Brazdil. "Introduction to the special issue on meta-learning." Machine learning 54, no. 3 (2004): 187-193.
[7] Thrun, Sebastian, and Tom M. Mitchell. Learning one more thing. CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE, 1994.
[8] Barto, Andrew G., Steven J. Bradtke, and Satinder P. Singh. "Learning to act using real-time dynamic programming." Artificial intelligence 72, no. 1-2 (1995): 81-138.
[9] Brazdil, Pavel B., Carlos Soares, and Joaquim Pinto Da Costa. "Ranking learning algorithms: Using IBL and meta-learning on accuracy and time results." Machine Learning 50, no. 3 (2003): 251-277.
[10] Eliasmith, Chris. How to build a brain: A neural architecture for biological cognition. Oxford University Press, 2013.
[11] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
[12] Yu, Yong, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. "A review of recurrent neural networks: LSTM cells and network architectures." Neural computation 31, no. 7 (2019): 1235-1270.
[13] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9, no. 8 (1997): 1735-1780.
[14] Caporale, Natalia, and Yang Dan. "Spike timing-dependent plasticity: a Hebbian learning rule." Annu. Rev. Neurosci. 31 (2008): 25-46.
[15] Miconi, Thomas, Kenneth Stanley, and Jeff Clune. "Differentiable plasticity: training plastic neural networks with backpropagation." In International Conference on Machine Learning, pp. 3559-3568. PMLR, 2018.
[16] Miconi, Thomas, et al. "Backpropamine: meta-training self-modifying neural networks with gradient descent." Workshop on Neural Information Processing Systems. 2018.