Course: Brain Inspired AI **Prof. Yalda Mohsenzade**
Name: Tina Gholami (251190343) **Assignment 3**

Paper: Back Propagation in the Brain

This paper studies the backpropagation procedure in the brain. In the cortex, synapses are all embedded within multi-layered networks, thus, it is hard to analyze an individual synaptic modification during backpropagation. Using deep neural networks, with the advents in the fields of AI and neuroscience, it is possible to evaluate brain's backpropagation process. The backpropagation algorithm learns quickly by computing synaptic updates using feedback connections to deliver error signals. Although such feedback connections are present in many parts of the cortex area, but still it is hard to formulate how it exactly works. But as suggested by the paper and according to the backpropagation algorithm, feedback connections are quite distinct for different neurons and synaptic activities, therefore such differences will act as error signals and hence yields effective learning process in the brain. However, there are also problems regarding the backpropagation such as their computations, demanding synaptic symmetry in the forward and backward paths, error signals becoming signed and extreme-valued, and feedback altering neural activity. Although it is mentioned in the paper that evolution and also brain's power has significantly affected the whole process, but still, there is no definite founding that contradicts the notion of brain recognizing differences in synaptic activities (aka, error signals) in order to employ backpropagation algorithm. Other theories have also been suggested in order to understand how the brain does the learning process, such as no feedback networks, scalar feedback networks, and node perturbation methods, but none of the above works as good as backpropagation to model the synaptic activity in brain's network. Next NGRAD hypothesis and implementation has been suggested to best show how brain might be actually exploiting backpropagation which is based on the idea of using neural activity differences to encode errors. NGRADs resolve significant implausibilities of backpropagation in a way that is intuitive and consistent with how we think biological circuits operate. They do away with the explicit propagation of error derivatives and instead compute them locally through differences in propagated activities and even areas within the synapses. There are still missing parts to this story but based on all the research conducted so far, it seems likely that a slow evolution of the thousands of genes that control the brain would favor getting as close as possible to computing the gradients that are needed for efficient learning of the trillions of synapses it contains.

Paper: Feedback Alignment in CNNs

Similarities between neural representations in biological networks and in deep artificial neural networks has lead researchers to develop analogies between the backpropagation learning algorithm used to train artificial networks and the synaptic plasticity rules operative in the brain. But there are challenges with the biological plausibility of implementing backpropagation. One is the symmetric forward and backward synaptic weights (weight transport problem). A number of methods have been proposed that do not rely on weight symmetry but, so far, these have failed to scale to deep convolutional networks and complex data. Modification of the feedback alignment method that enforces a weaker form of weight symmetry is suggested which requires agreement of weight sign but not magnitude in a way that can achieve performance competitive with backpropagation. There are related works done on such implementations, including MNIST handwritten digit dataset, CNNs, implementing Direct Feedback Alignment (DFA), CIFAR-10 dataset and ImageNet related architectures, and also a recently published work by Xiao et al. (2018b). Next, feedback alignment method is exhaustively studied with the according modifications to improve its performance on deeper neural networks, such as deep CNNs. These modifications include controlling the magnitude of the error signals that involves Initialization method, Excitatory/Inhibitory (E/I) method to freeze the sign of the forward weights after a certain number of epochs of training, and also the strict normalization (SN) method. The results show that the feedback alignment methods are competitive with, and in some cases exceed, backpropagation performance. Next, the paper tries to scale to deeper networks. But there are obstacles regrading this such as vanishing and exploding activations and error signals (the solution is the careful initialization of the forward and backward weights and also explicitly managing the size of the weights), Gradient signal alignment (Solution: USF method is applied but still in the deeper networks the performance is not good and the alignment angle does not drop), suggesting modifications to BP (meaning trying to train the model using BP but with modifications so that it mimic FA, and the solution is to add noise or to add L2 penalty to the loss function), Constraining the magnitude of the weights (which is using a variance-preserving initialization method, and thus, there will be no need for additional mechanisms to normalize the backward weights as the forward weights change), and effect of constrained weight signs (it is known that Freezing weight signs has a significant negative impact on learning but it depends on the weights stability).