The learning capacity of a convolutional neural network can be intuitively defined as the complement of the amount of learning deficiency that exists in the neural network at convergence. If a neural network reaches a suboptimal minima at convergence, then the learning capacity of the network is not utilized to its fullest potential. The term learning capacity can also be used in the context of a pool of neural networks. If all the networks in a pool have identical weights, then the total knowledge of the pool is no more than the knowledge acquired by any one single network in the pool. This also points to an underutilization of learning capacity. In this thesis, we explore different ways to push networks to maximally utilize their learning capacities in both an intra-network setting (within a single network) and in an inter-network setting (pool of several neural networks). For enhancing intra-network learning capacity, we propose a backward pass gradient modification method called PowerGrad Transform, in which we carefully modify the gradients before update to force the network to explore regions of the loss landscape that are inaccessible to standard methods of training. We find networks trained with the PowerGrad Transform method produces significantly higher training and test accuracy metrics, leading to an improvement in learning capacity. In the inter-network scenario, we use Feature Difference Loss functions to map shared feature representations from multiple networks in the pool to a single real number. We then invoke adversarial training to optimize the negative of the feature difference losses and force networks in a pool to produce diverse feature sets.