Assignment 2 - Group: (write your group number)

Tutors: (list all your tutors)

Group members: Firstname1 Lastname1 (unikey1), Firstname2 Lastname2 (unikey2), Firstname3 Lastname3 (unikey3)

Abstract

Abstract text goes here, justified and in italics. The abstract would normally be one paragraph long.

Introduction

This template should be used as a starting point for your report.

Previous Work

**Methods Used**

Image Classifications are complex tasks that require extraction of features and key attributes from images before they can be pumped into a model for classification. Often times, these feature extraction processes impose a massive challenge as images come in all manners of sizes (dimensions), lighting conditions, intensities of pixels, angles, scale etc. This makes it difficult for a computer to understand the key features before a classification can be carried out.

Although in humans understanding & processing images comes as a second nature, the in the field of computer vision the computer simply gets a large array of numbers to process. The figure below depicts a clear picture of this challenge.

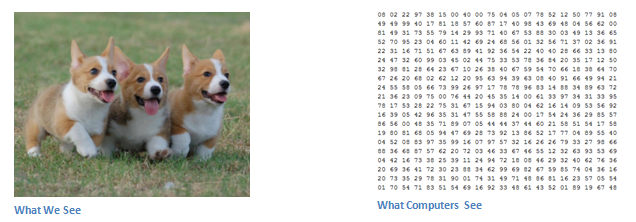


Figure 1: Computer vision vs our vision (<https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/>)

Our chosen dataset (Cifar10) is such a problem where we have thousands of 32x32 images containing 10 distinct classes of objects. We have opted to approach this classification task from a chronological technology/technique/method in the field of computer vision.

Below is a list of methods that we have attempted in chronological order

1. **Sift/Daisy Feature Extraction Followed by Classification**

Before the advent of more modern techniques such as convolutional neural networks the field of computer vision utilized several scale invariant algorithms (Sift, Surf) to carry out manual feature curation/extraction before pumping the said features into a classification algorithm. Although a lot of algorithms have been developed to extract features this is still a manual, unreliable, and inconstant methodology that hardly yielded good performance.

1. **Convolutional Neural Network (CNN)**

A more modern approach in the field of computer vision where we utilize an artificial neural network alongside a technique/methodology of convolutions to understand features and predict the image classes after passing it through a multi-layer neural network. This method yields a high degree of accuracy and able to understand more complex distributions of the data.

1. **Recurrent Neural Network (RNN – LSTM & GRU)**
2. **Residual Neural Network (ResNet)**

Another variation of a multi-layer neural network that still utilizes convolutions in the heart but with a slight twist called “short-cuts” that addresses the vanishing gradient problem in a deep neural network and thereby allows build and construction of deeper networks. ResNets are able to understand even more complex features simply via the convent of a deeper network and ability to train deeper for longer.

**Background on SIFT (Scale Invariant Feature Transformation)**

There are several techniques to extract features from images such has Histogram of Gradients (HoG), Binarizing and blurring, corner detection (Corner Harris and corner peak) but by far the most widely used method is SIFT keypoint detector.

Sift is effective due to the fact that it is able to detect/match features between images even if the scale, orientation, viewpoint, and illumination are different between images. The SIFT algorithm takes a grayscale image and generates interest points (keypoints) from the image where the local gradient orientation histograms of the image intensities are collected and statistically summarized to produce a keypoint descriptor of the local image structure (Prof. Tony Lindeberg, 2012, Scholarpedia, 7-5:10491). Typically, these statistics are gathered from a surrounding neighborhood of each keypoint.

We have opted to utilize a variation of SIFT (since it is a proprietary algorithm) called “Daisy feature extractor” which is available in the scikit image library. Once these descriptors from each image captured they can be utilized for image classification tasks as well as image matching.

**Background on Multi-Layer Neural Network**

The key advantage of a multi-layer neural net is that is able to predict and model on any distribution of data and able to create non-linear decision boundaries. The figure below depicts a simple artificial neural net and also a multilayer neural net.

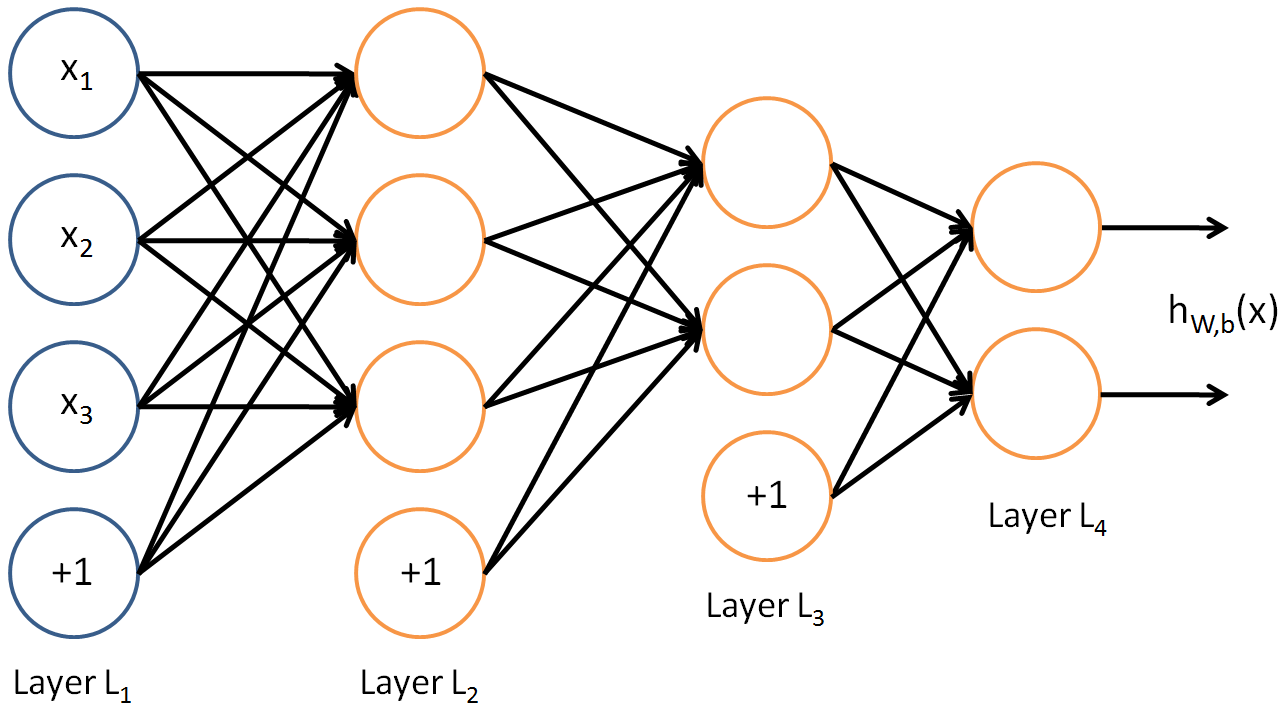
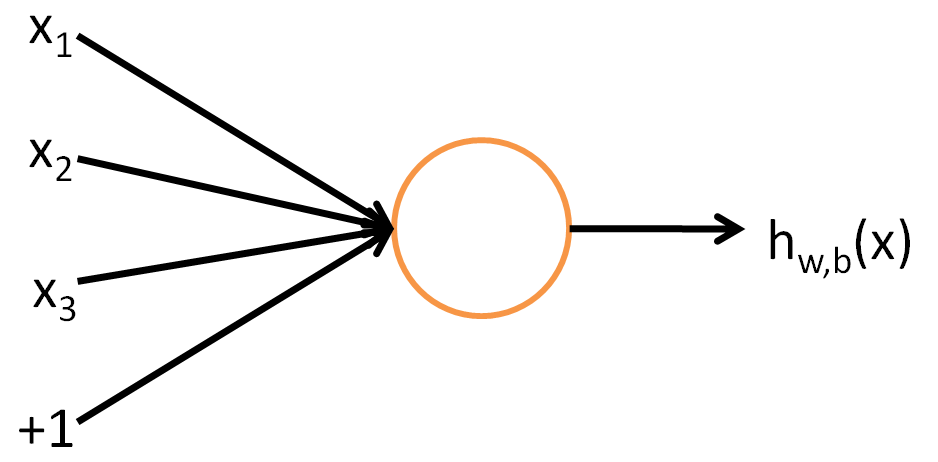
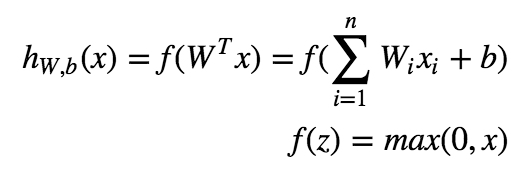


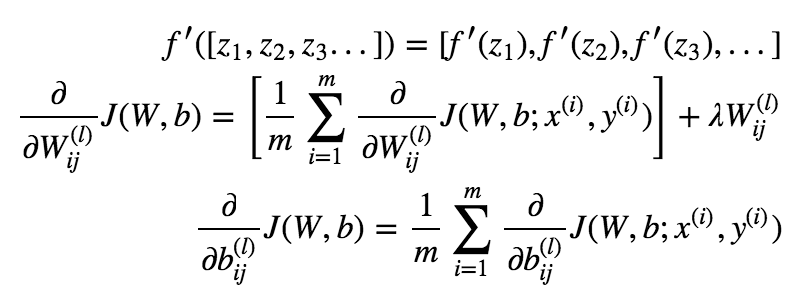
Figure 2: Stanford University depiction of a generic feedforward multi-layer neural network.

The key concepts of a neural nets are the inputs, neurons, the weights and bias terms, the output/activation function, and the optimization function (gradient descent/backpropagation). Each circle depicted in the image above is a single neuron or a computation unit which takes in the input (x1, x2, x3, and the bias unit) and outputs some new x via an activation function. The output from one neuron is then passed onto (feed forward) to the next layer of neurons which in turn carries out a similar exercise of applying an activation function.

There are several activation functions that are in practice but for our exercise we have opted to utilize a Rectified Linear Function (ReLU) in all our cases. Recent research suggests that ReLU activations perform better in deep neural networks when compared to its counterparts (UFLDF Tutorial on multilayer neural network – Stanford University). The formula that runs through each of the neurons in our case can be summarized below (where n is the number of x inputs and f is the activation function) -

****

Optimizations in a neural network is carried out with gradient descent however, there are several layers within a deep neural network therefore we utilized backpropagation to optimize the weights within the network. Backpropagation simply put, allows the easy calculation of the partial derivatives of the cost functions for each layer. The chain rule allows easy calculation of the derivative of the overall cost function ((UFLDF Tutorial on multilayer neural network – Stanford University). The formula of the overall cost function for the network is outlined below.



**Convolutional Neural Network (CNN)**

CNNs are by far the most popular method for image classification tasks. The core idea of CNN are the convolutions which learns various key features within an image and thereby able to utilize them for image classifications and identification. In generalized terms, convolutions allow the network to learn edges, orientations, colors, blotches, blobs and allows neurons to activate when similar edges, orientations etc. are identified within another image.

Before CNNs are can be full explained a brief understanding of a regular multilayer neural network is required.

Image classifications are complex tasks a

**Residual Neural Network (ResNet)**

**Sift/Daisy Features & SVM**

**Experiments**

**Conclusions**

**References**

1. Pryor TA, Gardner RM, Clayton RD, Warner HR. The HELP system. J Med Sys. 1983;7:87-101.
2. Gardner RM, Golubjatnikov OK, Laub RM, Jacobson JT, Evans RS. Computer-critiqued blood ordering using the HELP system. Comput Biomed Res 1990;23:514-28.