Neural Networks : Inspiration in Biology

compiled by Alvin Wan from Professor Jitendra Malik's lecture

We will preface this note with insights from biology.

1 The Brain

Let us consider the brain; we can see it as a sheet of processors. Each lobe in the brain is subsumed by some activation function. For example, the temporal lobe is associated with visual processing, and the parietal lobe is associated with motor control. The frontal lobe is concerned with higher-level planning. 30% of your body's energy is consumed by the brain. What consumes so much energy?

There are 10^{11} neurons, where each has about 10^4 connections. This makes 10^{15} weights.

Let us now consider one of these neurons. Signals are transmitted across the axon, and small junctures - called synapses - exist at the near-intersection of the receiving neuron's dendrite and the sender's axon. If the neuron is not getting input, then this neuron does not activate. If the neuron is, it will experience approximately 200-300 "spikes" a second. This is several kHz. Our computers can run at several Ghz, but your brain contains many more neurons, allowing for more parallel processing.

We now liken this to our neural networks, where each node is considered a neuron $x_i^{(l)}$. We take **synaptic weights** between each neuron, w_{ij}. We represent the activation of a neuron using the activation function g.

2 Eye

The photoreceptors, which are like pixels, are divided into various layers: ganglion cells, bipolar cells, horizontal cells, rods, and cones. We can think of each of these as layers in our neural networks.

Start with input: neurons that respond to light, rods and cones. These form your pixels, or the x_i. They are organized in a 2-d sheet and pass along information through several layers. The output of the ganglion cells is what is called an optic nerve, which connects the eye to the brain. There is some interesting phenomenon happening with the bipolar cells. Each bipolar cell gets direct - **"center"** - input from multiple photoreceptors. Note that the horizontal cell also connects photoreceptors *and* bipolar cells, so bipolar cells receive indirect - **"surround"** - input from photoreceptors.

We call positive w_{ij} **excitatory** input and negative w_{ij} would be **inhibitory**.

We can use two neurons to represent a negative quantity. We simply have two units where one fires with positive input and the other with negative input.

What are the takeaways? Remember "synapse", "excitatory", and "inhibitory".

2.1 Receptive Fields

We consider the **receptive field** of a neuron to be the pixels that affect the output of that neuron. The higher the layer, the larger the receptive field. The receptive field of a retinal ganglion cell can be modeled as a "Difference of Gaussians", as this embodies the notion of excitatory center input and inhibitory surround input.

There is only local connectivity; nodes in each layer are not connected to all nodes in the next layer.

3 Visual Processing

We have various visual processing areas, named $V1, V2\dots$ in our brain. Likewise, those that control movement are $M1, M2\dots$ (motor control), and those that control hearing are $A1, A2\dots$ (auditory). We note a phenomenon called **orientation selectivity** in $V1$. Our experiment shows that a line of a certain thickness at a certain angle in a certain position will trigger $V1$. This makes sense, as each neuron only has local connectivity.

In this experiment in 1962, Hubel and Wiesel discovered that "complex cells" could detect the same pattern across different positions. This is the motivation for our convolution kernels, or filters.

We focus on the flow of information in the brain, called the **ventral stream**. Starting out, the receptive field is fairly small. We find that successive layers increase their receptive fields, and this is the motivation for multiple layers of convolutions.

Our interlude into biology ends here, and we continue with early models of neural networks in the next note.