Conor Lorsung

Dr. Heroux

CSCI-Capstone

2/3/2018

Neural Networks and Image Processing

Neural networks sound like they could come right out of a sci-fi film. However, they have been around us for quite some time and continue to benefit us today. Beginning with simple neuron models and rule sets built to identify basic shapes, they have been pushed to help us in areas ranging from geology to cognitive science. In addition to this, they have been taught to teach themselves using algorithms like backpropagation and hill-climbing. In the future, they could very well make us music and tell us about the intricacies in our genetic makeup. To understand how this may be possible, it is necessary to cover the history, technical underpinnings, and current state of the field of neural networks.

Visual perception, specifically related to saliency and pattern recognition, is a complicated subject. In the 1950's, D. O. Hebb studied rats and their visual acuity by rearing them in darkness, then exposing them to light and images for a total of 15 minutes. It was then shown that despite having very limited visual experience, the rats could distinguish between patterns and objects. [1] Operating under this premise, it can be said that there exists a system or set of rules that can be used to learn to see. Thus, it can be applied to computation. It is not as easy as telling a computer to look somewhere in a picture and choose an object. First, the computer needs to be able to register input.

The Perceptron was one of the first of its kind. It used a novel network of neuron-type circuits to determine if it has been given a square or a circle. The subsystems themselves take various inputs from photo-sensors and examined the voltages. [2] If a specified threshold had been exceeded, the neuron (circuit) would fire and pass along an impulse. This allowed the Perceptron to determine where there was light coming through, and where the image was opaque. In a more general sense, the neuron model operates by taking various inputs, weighting them, and then passing them into the threshold evaluator which then may pass a value onto a result aggregator. This was the predecessor to the image-processing neural networks we see today. Alongside the development of new machines to analyze images, algorithms were being designed to optimize these systems. Marvin Minsky details the *hill-climbing* heuristic that allows neural networks to come to conclusions using their thresholds and outputs. Hill-climbing explains that for a given set of inputs $\lambda_i \ldots \lambda_n$ and an output $E(\lambda_i \ldots \lambda_n)$, assuming we want to maximize the output E, it will modify the input to find maxima for E. [3] This heuristic is the foundation for some of the image processing techniques used today in neural networks.

As the field has progressed, it has become more complex. This appears in more robust neural networks, better techniques for processing input, and solutions to problems such as learning. The neuron model mentioned before is the technical foundation of every neural network and each of the above sections. In a lecture given by Patrick Winston, the neuron model is explained in depth. It is composed of $m+1$ inputs corresponding to input signals $x_1 \ldots x_m$ that are processed through weights $w_1 \ldots w_m$ and given to a threshold t. [4] What this means with regards to image processing is that the stronger a color may be, or the more contrasting a color may be at a given point will influence the final calculation at that neuron, pass the threshold, and trigger a pass-forward of information (synapse). Once the neuron model is understood, we should look at our data. If we look at an image with vivid colors, hue may be interpreted as intensity. This is not always the case, however. Therefore it is necessary to normalize the data.

L. Itti, et al. demonstrates a technique for separating hue from intensity:

> … an intensity image I is obtained as $I = (r + g + b)/3$. I is used to create a Gaussian pyramid $I(\sigma)$, where $\sigma \in [0..8]$ is the scale. However, because hue variations are not perceivable at very low luminance (and hence are not salient), normalization is only applied at the locations where I is larger than 1/10 of its maximum over the entire image (other locations yield zero r, g and b). [5]

This demonstrates how the colors are normalized in this example and lays out the idea behind normalization for further cases of hue and intensity separation. The normalization process (if set up correctly for the data set) will ensure that the values contained in our data set will fit our model. This is a crucial part of developing a working system. Once the data is normalized, it can be fed into the neural network. At this point, nothing will be evaluated correctly since no weights have been designated. One integral algorithm for developing weights is called *backpropagation*. It is crucial to the neural networks learning ability.

Backpropagation refers to calculating the weights for each node in the neural network with a given training set. It works by taking the partial derivative of a cost function with respect to any weights w or biases b in the network. [6] This is written as $\partial C/\partial w$ and $\partial C/\partial b$ respectively. Nielsen provides an equation $(\partial C) / (\partial w^l_{jk}) = a_k^{l-1} \delta^l_j$ that gives the rate of change of the cost equation respecting any one weight, where C is our cost equation, w is the weight, a is the activation of the neuron, and δ is the error of the output. [6] Knowing the rate of change of the cost gives the algorithm enough information to generate learned weights. At this point, backpropagation will teach weights if the input neuron is low-activation, or if the output neuron is either high-activation or low-activation. [6] Using the above information, I will develop a prototype that selects the salient areas of an image.

The prototype I will develop will use the method described by Itti et. al. to determine the salient locations of an image. Ideally, it will be robust to noise. To do this, I will attempt to implement a neural network in Python. Following the outlined backpropagation algorithm provided by Nielsen (which inherently uses gradient descent for optimization), the neural network will implement this process so it may learn from test images. The core structure of the prototype will contain an array of artificial neurons with hand-selected beginning weight values. This will allow for backpropagation to teach the network which weights work and which do not. In effect, I hope to do relatively little work hand-selecting the values for the weights. In addition to the backpropagation algorithm, I would like to implement Minsky's hill-climbing algorithm as described above. Being able to compare and contrast the hill-climbing algorithm with gradient descent and backpropagation would be interesting. I would like to find if there is a significant computational time difference between the two and if one yields more accurate output than the other. It may also prove interesting to check the prototype against human eyesight to see how accurate the model is.

The current state of the field of neural networks is broad. Neural networks have been implemented for many different purposes in a variety of disciplines. Image recognition, medicine, geological sciences, and law enforcement all use these systems in many diverse ways. First, we have already examined a little bit how neural networks can be used to predict the salient areas of an image. [5] Going more into depth here with this, image saliency can be used to model eye behavior. Moreover, the model developed by Itti et. al. is robust to noise as shown in figures 4 and 5 in the article. This system can be used for data visualization enhancement, text detection, and has uses in cognitive science. Law enforcement has also found uses for neural networks and image processing. License plate scanners can use trained neural networks paired with a multithresholding technique to quickly and accurately read plates as detailed by Bhushan et. al.. They developed a system that was able to perform with an accurate recognition rate of 98.40%. [7] It works by capturing an image of a license plate and rasterizing the image. The colors can be normalized at this point using thresholding techniques that enhance letters on backgrounds. From there, text analysis using a neural network identifies the letters and numbers on the plate. This can aid in the efficiency of red-light cameras, speed cameras, and vehicle-bound monitoring equipment.

Another area where neural networks and image processing are used is in medicine. Currently, there are not too many medical images that can be used to train neural networks to find landmarks and abnormalities. Additionally, the complexity of 3D medical images is much greater than that of a regular 2D image. Therefore, there will be many more weights that need to be calculated if an accurate model is expected. The task then became, instead of creating more images, to train convolutional neural networks (CNN) using new techniques. Zhang et. al. developed a deep learning approach that tackled this problem. Using their two-stage task-oriented deep learning (T^2DL) model they were able to achieve a mean error of 2.96mm in brain landmark detection using 1200 landmarks and 700 subjects. In prostate landmark detection with 7 landmarks and 73 subjects, they were able to achieve a mean error of 3.34mm. [8] Where doctors would have to parse each picture and mark down points of notice beforehand, these systems can do the job for them maintaining accuracy and speed. The aforementioned CNN is able to detect thousands of landmarks simultaneously in around 1 second. [8] These recent techniques are impressive, and lead the way to more novel developments.

Neural networks have a promising future ahead of them in various fields. Geology, medicine, and music to name a few, are areas that will continue to benefit from unique implementations of these systems. The work is being laid out for researchers to improve upon their models as Maier et. al. has shown. Their research focuses on categorizing and examining current artificial neural network (ANN) implementations that attempt to predict water resource variables in river systems. They begin by developing taxonomies of approaches that they found in 210 journals published between 1999 and 2007. From this, ANN models can be outlined and developed. After studying the journals and categorizing them, it was found that there is not enough research on the methodologies of these systems and ensuring that they work properly. Most of the models were found to be ad-hoc. In addition to this, they found that many journals used a linear approach to determine input significance, which is counterintuitive to the non-linear nature of ANN's. [9] By researching the areas that are due for improvement, ANN's and other NN systems may be further improved in the future for more accurate prediction models.

Building on the neural network models present today will lead to an improvement in medicine. Eric Roberts, a professor at Stanford University envisions a bright future for neural networks and their application towards analyzing the human genome. [10] It may be that they can help us discover trends that went previously unnoticed, or we simply could not see. By implementing introspective research like the work of Maier on river resources, medicine could see a large impact from improving methodologies when developing CNN's or ANN's.

Music is another area that is currently being researched with an interest of applying neural networks. Daniel Johnson developed a recurrent neural network (RNN) that employs a feedforward algorithm to learn. It operates by having each node (neuron) sum its inputs, multiplying them by a weight, and then adding a bias value. The numbers are then normalized to a predefined range with a nonlinear function. An example of this type of function would be a sigmoid function. His model has many layers of neurons in it, which allows it to better learn with a provided set of data. [11] This particular approach uses backpropagation to learn. Instead of simply building a neural network and feeding music into it, he developed his own properties to better suit the task of making music. Networks like these are being developed to push technology into fields it might not immediately be thought of having a role in. Eric Roberts also says that someday neural networks will be able to recreate compositions similar to those of Mozart.

The neural networks of today are built on the foundations of seemingly simple machines that were assigned tasks like determining if a slide had a circle or a square on it. Today, with the development and continuous research of algorithms such as backpropagation, forwardpropagation, and hill-climbing, neural networks are able to find anomalies in medical images, predict water resources, and accurately model the human eye. Where their learning used to be supervised, it can now be automated. Neural networks today contain hundreds, if not thousands of artificial neurons and seem to be increasing in size. As with the leaps we have seen from squares and circles to landmark identification in brain scans, we can nearly expect to see the same type of growth for the future. Researching the methodologies and developing better models could help us in the future make new genres of music, tell us something about our human construction, and maintain an ecological balance with the world around us.

Bibliography

1. D.O. Hebb, *The organization of behavior : a neuropsychological theory*, John Wiley, 1964.

2. F. Rosenblatt, *The perceptron, a perceiving and recognizing automaton <Project Para>*, Cornell Aeronautical Laboratory, 1957.

3. M. Minsky, “Steps toward Artificial Intelligence,” *Proceedings of the IRE*, vol. 49, no. 1, 1961.

4. P. Winston, *Lecture 12a: Neural Nets*, Massachusetts Institute of Technology, 2015.

5. L. Itti, et al., “A model of saliency-based visual attention for rapid scene analysis,” *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, 1998, pp. 1254-1259; DOI 10.1109/34.730558.

6. M. Nielsen, “How the backpropagation algorithm works,” *Book How the backpropagation algorithm works*, Series How the backpropagation algorithm works, ed., Editor ed.^eds., Determination Press, 2015, pp.

7. B. Bhushan, et al., *License Plate Recognition System using Neural Networks and Multithresholding Technique*, 2013, p. 45-50.

8. J. Zhang, et al., “Detecting Anatomical Landmarks From Limited Medical Imaging Data Using Two-Stage Task-Oriented Deep Neural Networks,” *IEEE Transactions on Image Processing*, vol. 26, no. 10, 2017, pp. 4753-4764; DOI 10.1109/TIP.2017.2721106.

9. H. Maier, et al., *Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions*, 2010, p. 891-909.

10. E. Roberts, “Future of Neural Nets,” https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Future/index.html.

11. D. Johnson, “Composing Music With Recurrent Neural Networks,” 2015; http://www.hexahedria.com/2015/08/03/composing-music-with-recurrent-neural-networks/.