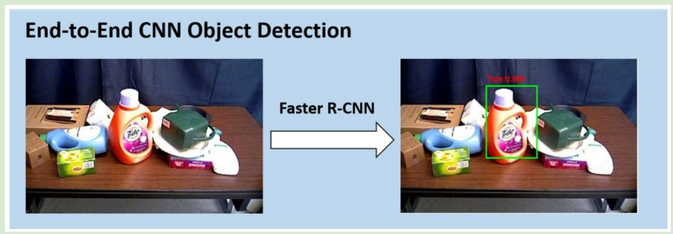
Implementing And Evaluating SqueezeNet

Caitlin McElwee, Iris Bahar
School of Engineering, Brown University

Lab's Previous Work

- Convolutional Neural Networks (CNNs) can be used to draw bounding boxes around objects in images for object detection. This is important for scene perception in robotics.



- Multiple network architectures have been evaluated for use.

Research Objective

- Implement the squeezenet CNN architecture with the lab's previous work, evaluate its performance against current methods.

What is a Neural Network?

- A neural network is a mathematical model organized in layers of neurons: small cells that each take information in, modifies it in some way, then pass its output along to the next layer.

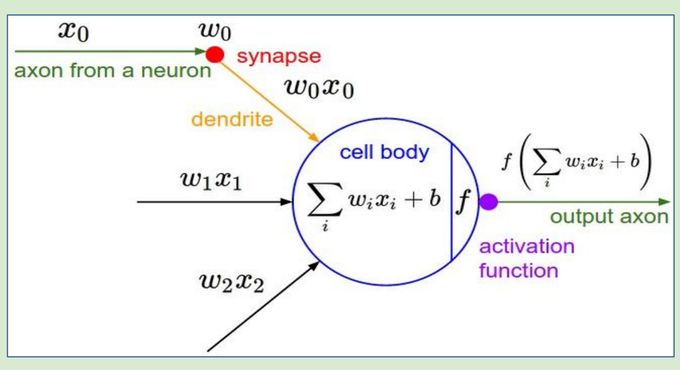


Figure 2: Each neuron in the network takes in weighted input, performs some operation, then outputs to the next layer. Source: https://i.stack.imgur.com/7mTvt.jpg

- Having more neurons allows for more connections, which allows for greater precision for complex tasks, but requires more training and computation to reach peak accuracy.

What is SqueezeNet?

- Every neural network has its **architecture** - how the network is laid out.
- Standard architectures include AlexNet, VGG16, Res50.
- The above-listed architectures are all large,all requiring 500MB or more to store a model.
- SqueezeNet is small - taking up only 10MB (50x more space-efficient), with accuracy rates comparable to AlexNet.
- Much smaller size allows installation on devices that cannot hold larger networks, such as phones or embedded circuits.

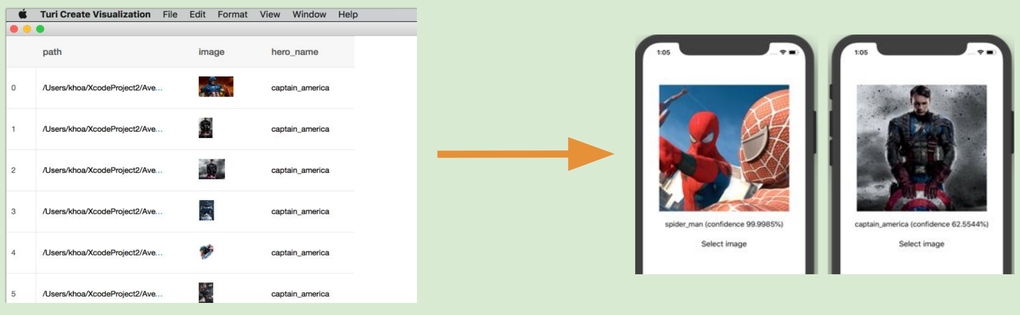


Fig 3: Tuli, an iOS app, uses SqueezeNet to perform image classification on the phone instead of in the cloud, which saves time from connection latency and money from server use costs. Sources: https://cdn-images-1.medium.com/max/1600/1*37FVIMCxy_hNmyJjUonWZA.png https://cdn-images-1.medium.com/max/1600/1*GFVX6a9oESRFzNHNjpLUvw.png

Equipment

- Ubuntu 16.04 Desktop
- nVidia GTX Titan Xp graphics card

AlexNet vs SqueezeNet

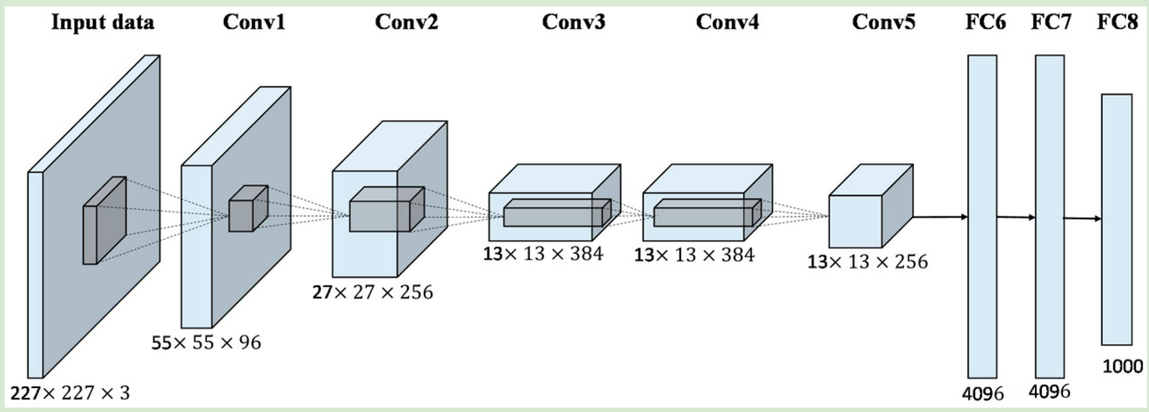


Figure 4: AlexNet's architecture is presented, with a depth column highlighted. Source: http://www.mdpi.com/2072-4292/9/8/848

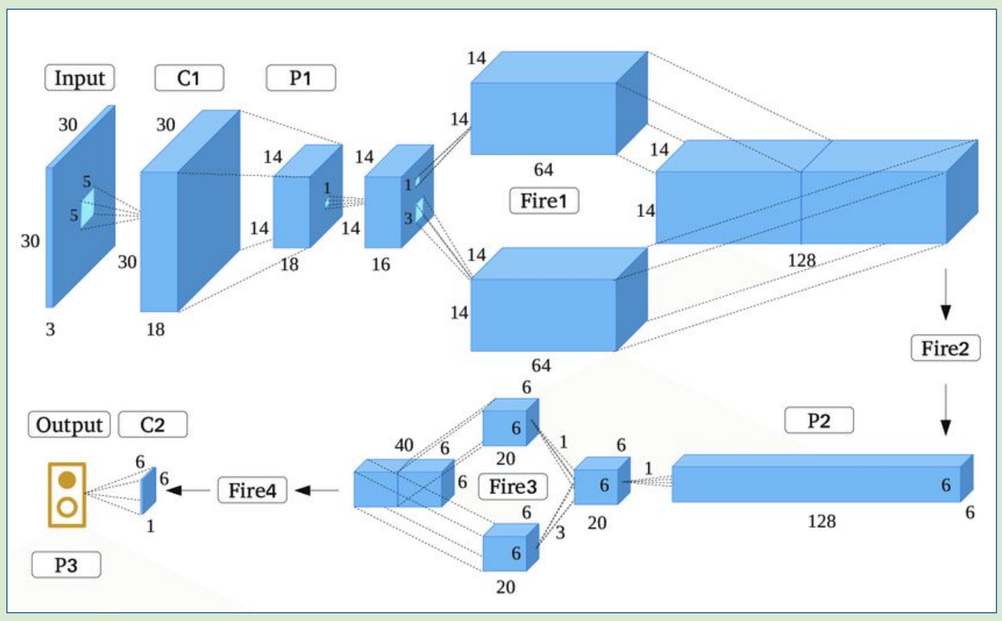


Figure 5: SqueezeNet's architecture is presented, with some repetitive sections skipped for space. Source: https://www.researchgate.net/figure/SqueezeNet-like-architecture_fig8_320723863

- As can be seen in the above figures, Alexnet has three FC (fully-connected) layers while SqueezeNet has none. FC layers are the most space-consuming kind of layer, so by having none SqueezeNet is able to be much smaller.

Potential Leads

- Pytorch, a python library for neural networks; the lab's current method.
- Caffe, a similar python library; known for its collection of pretrained models.
- Pytorch again, using new methods and dataset.
- Training a model from scratch on ImageNet, an iconic massive dataset (~500,000 labeled images).

Results

- Caffe, due to difficulty of setting up and redundancy with Pytorch, was abandoned.
- Pytorch proved most promising library; its pretrained models lackluster for our purposes.
- Pretraining a model on ImageNet most potentially successful lead.
- Pretraining still running as of now.

Future Work

- Implement fully-pretrained model into lab's training and testing suite.
- Evaluate ImageNet-pretrained model of SqueezeNet against AlexNet, VGG16, Res50, etc.