Jetson Nano and Transfer Learning

Edge devices have revolutionized the world of machine learning and AI. Fifteen years ago, no one would have been able to predict that it would be as simple and easy to deploy machine learning models on devices that could fit in the palm of your hand. Currently the Jetson Nano is one of the most versatile edge devices on the market, with over 4 GB of RAM and fitted with a 128-core Maxwell GPU, it allows for deployment of computer vision modules with as few as 10 lines of code. Transfer learning has also propelled the field forward, allowing for AI enthusiasts everywhere to harness the power of models trained on terabytes of data (Shavilk. AI. ,2010).

My project focused on the deployment of an object detection model on the Jetson Nano. The underlying model that I used was the Mobilenet model which is a transfer learning framework focused on mobile vision applications (Howard et AI. ,2017). The Mobilenet architecture leverages depthwise separable convolutions in order to decrease the amount of computations needed to train and deploy models. One of the strongest features of Mobilenet is the ability to decrease computations and model size by applying a single filter to each input channel, then the pointwise convolution a 1 by 1 convolution to join the outputs of the depthwise convolution in a single step. This factorization allows for reduction of complexity in the model which makes it a great framework to build on top of.

$$= \frac{D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F}{D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F} = \frac{1}{N} + \frac{1}{D_K^2}$$

Figure 1.1: Reduction of Complexity VIA DSC

When DK * DK is 9, the reduction in complexity is over 8 times. This makes a huge difference when dealing with an edge device which only contains 1-4GB of RAM. Although there is a huge reduction in complexity ,the difference in accuracy is negligible when compared to more computationally expensive methods.

Table 8. MobileNet Comparison to Popular Models

Model	ImageNet Accuracy	Million Mult-Adds	Million Parameters
1.0 MobileNet-224	70.6%	569	4.2
GoogleNet	69.8%	1550	6.8
VGG 16	71.5%	15300	138

When compared to GoogleNet and VGG , MobileNet has approximately the same amount of accuracy ;however, there is a huge reduction in the amount of parameters. This allows for quicker and more efficiency training without the large penalty of accuracy. Overall, transfer learning methods fill in the lack of computational power when working with edge devices. Moving forward, it will be intriguing to see how various cloud computing companies work with edge devices in order to make the deployment process even more simple.

Sources:

Torrey, Lisa, and Jude Shavlik. "Transfer learning." *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI global, 2010. 242-264.

Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).