Implementation of Pet Detection and Recognition on the NVIDIA Jetson Nano

1st Chestnut, William
The Governor's School for Science and Technology
MITRE
Hampton, United States
william.chestnut@nhrec.org

2nd Cox, Jason
The Governor's School for Science and Technology
MITRE
Hampton, United States
jason.cox@nhrec.org

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

Machine learning and artificial intelligence are both becoming a larger part of today's world with seemingly endless applications in all sectors such agriculture, business, and many more. While it has great use cases and problem solving capabilities, machine learning requires a considerable amount of computing power with some algorithms pushing even the most high level GPUs to their limits. This need for high computational power is not always feasible, especially in situations where low power consumption is required. To help deal with this demand for low power options for machine learning, companies like NVIDIA and Raspberry Pi have created their line up of embedded heterogeneous systems such as the NVIDIA Jetson series and Raspberry Pi series. Embedded heterogeneous systems, single motherboard computers that have components with distinct parts, such as the Jetson Nano and Raspberry Pi 4 have the goal of providing decent performance at a very low power consumption [1]. By doing this and designing specifically for machine learning, as the Jetson line up has, these embedded platforms create a low power and portable option for machine learning that could become a very important part in the future of artificial intelligence as a whole [5].

II. Background and Overview

A. NVIDIA Jetson Nano Performance

Heterogeneous platforms (which contain a CPU and GPU) are crucial for low power machine learning. Vision based machine learning deals with matrix data structures, which the GPU, with its thousands of core processors, is optimized to work with. Machine learning and image processing consists of extremely parallelizable operations that run much faster on GPUs compared to CPUs, which is why a heterogeneous architecture is crucial for machine learning performance.

The three main platforms utilized in research regarding embedded machine learning include: the Raspberry Pi 4, the NVIDIA Jetson Nano, and the NVIDIA Jetson TX2. As shown in the figure below, the Raspberry Pi 4 has the lowest cost but also has the lowest performance, while the NVIDIA Jetson TX2 provides the highest performance with by far the highest cost. Meanwhile, the NVIDIA Jetson Nano is priced around 50 dollars above the Raspberry Pi yet still provides a relatively large amount of performance. This gives the Jetson Nano the highest price to performance among the three cards and is part of the reason why many researchers use it as their platform of choice.

This adequate performance shows in benchmarks of machine learning algorithms where the NVIDIA Jetson is able to consistently achieve speeds of 15 FPS and higher on a multitude of object detection algorithms using Tensorflow [6]. Its performance for classification algorithms with Tensorflow are even higher at 64 and 36 FPS [6]. While these numbers may seem low considering the average screen runs at 60 FPS, they are more than adequate for the mobile machine learning use cases the Jetson Nano is designed for.

	Raspberry Pi 4	Jetson Nano	Jetson TX2
Performance	13.5 GFLOPS	472 GFLOPS	1.3 TFLOPS
CPU	Quad-core ARM Cortex-A72 64-bit (1.5 GHz)	Quad-Core ARM Cortex-A57 64-bit (1.42 GHz)	Quad-Core ARM Cortex-A57 (2 GHz) + Dual-Core NVIDIA Denver2 (2 GHz)
GPU	Broadcom Video Core VI (32-bit)	NVIDIA Maxwell w/ 128 CUDA core (921 MHz)	NVIDIA Pascal 256 CUDA cores (1300 MHz)
Power under load	2.56W - 7.30W	5W - 10W	7.5W - 15W
Price	$35	$89	$399

[1]

B. Leveraging Convolutional Neural Networks

One of the main tools used to improve image classification algorithms, especially with animals, is Convolutional Neural Networks [1]. A Convolutional Neural Network (CNN) reduces image size while helping identify patterns in the image. For example, a CNN may look for a triangle pattern similar to the top of a dog's ear, in which it would produce an image with high values mainly near the dog's ear in the picture given. This also creates the opportunity to use layers of CNNs that look for different patterns, which identifies more features, making the input data more useful and accuracy higher [1].

While Convolutional Neural Networks help greatly increase the performance of computer vision algorithms, they can also cause an increase in processing requirement with their multitude of layers. This is why it's important to choose an optimal resolution and layer count for the intended task. Reducing the resolution of each layer can greatly help with increasing performance because it reduces the total amount of data the neural network must analyze [4]. However, this may reduce the accuracy of the algorithm because some data may be lost, but this effect is rarely large. The other primary approach to increasing performance of a CNN is by reducing the amount of layers, which similar to reducing resolution, will increase performance while potentially decreasing accuracy [4].

C. Demonstrated Applications

There are many applications for the NVIDIA Jetson Nano, but most of the projects employing the Nano involve image recognition. For example, researchers in Hong-Kong wanted to make a way for a computer to translate Hong-Kong sign language so deaf people can better communicate with others. Since they wanted this device to be portable, the usual high powered computing hardware was not an option for this application, so the researchers chose to utilize the NVIDIA Jetson Nano [5]. Using a 3D CNN algorithm with a resolution input of 480 x 640 and a sampling rate of 30 fps, the researchers were able to achieve an accuracy of 93.3% with a total response time of 5.82 seconds [5]. This is a perfect application of the NVIDIA Jetson Nano and shows what doors it can open in the world of machine learning.

Researchers have also been exploring the option of utilizing the NVIDIA Jetson platform to drive autonomous robots with LiDAR and optical positioning tracking [2]. However, since this requires a lot of processing power, the researchers opted to utilize the larger NVIDIA Jetson TX2 rather than the Nano. They tested algorithms including: ORB-SLAM2, VISO2, RTAB-MAP, SPTAM, and ZED-VO, and briefly described how each algorithm works and how they are applied to the robot [2]. The researchers concluded that "ORB-SLAM2 and RTAB-MAP" provided the most performance, but they did acknowledge that it may be necessary to consider changing the source code of the algorithms to help increase speed on the Nvidia Jetson platform [2].

D. Future Research

There have been many projects involving NVIDIA Jetson platforms, which have all shown the exceptional performance of the devices, but they have also described room for improvement both with the software and hardware. In terms of software, many researchers recommend exploring the option of changing the source code of algorithms to perform better on the Jetson platform [2]. While this will require a good amount of effort, it could prove to be worth it with large gains in performance. In terms of hardware, many researchers have recommended circuit level improvements such as smaller feature sizes to increase integration density and non-volatile memories to achieve near-zero idle power [4]. Researchers have also recommended microarchitecture improvements such as closer integration between the CPU and GPU to reduce data-transfer overheads [4].

Other researchers have suggested more general changes such as making the process of developing software on the NVIDIA Jetson platform easier [3]. Currently, in order to optimize for the NVIDIA Jetson, developers have to rewrite large amounts of their application for each processor [3]. They also suggest enabling better dynamic frequency scaling to allow the Jetson to change frequency for applications to optimize performance and power consumption. Lastly, the Jetson should also improve the memory bandwidth at low frequencies, which has been found to bottleneck programs that run at lower operating frequencies [3].

These options may take a lot of work to implement, but their performance benefit may prove to be worth it. In future research, the practicality of these options should be explored and researchers should demonstrate whether their performance increase is worth the effort.

References

[1] A. A. Süzen, B. Duman and B. Şen, (2020). "Benchmark Analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN," Isparta University of Applied Sciences.

[2] Giubilato, R., Chiodini, S., Pertile, M., Debei, S. (2019). "An Evaluation of ROS-compatible Stereo Visual SLAM Methods on a nVidia Jetson TX2." Industrial Engineering Dept., University of Padova

[3] Halawa H., Abdelhafez H.A., Boktor A., Ripeanu M. (2017). "NVIDIA Jetson Platform Characterization." The University of British Columbia, Vancouver, Canada.

[4] Mattal, Sparsh. (2018). "A Survey on Optimized Implementation of Deep Learning Models on the NVIDIA Jetson Platform." Indian Institute of Technology Roorkee.

[5] Zhenxing Zhou, Yisiang Neo, King-Shan Lui, Vincent W.L. Tam, Edmund Y. Lam, and Ngai Wong. (2020). "A Portable Hong Kong Sign Language Translation Platform with Deep Learning and Jetson Nano." The University of Hong Kong.

[6] NVIDIA. (2021, January 5). Jetson Nano: Deep learning inference benchmarks. NVIDIA Developer. Retrieved November 24, 2021, from https://developer.nvidia.com/embedded/jetson-nano-dl-inference-benchmarks.