Comparative Analysis of Hardware Performance Across ARM and x86 Architectures Using Computer Vision

Team Members:

Johnny Acuna , Brian Garcia, Raymond Palacio, Brandon Siu

Email:{jjacuna,briang1,rapalacio,brandonsiu}@cpp.edu

Faculty Advisor: Professor: Mohamed El-Hadedy Aly

Email:mealy@cpp.edu

California State Polytechnic University, Pomona Department of Electrical and Computer Engineering, College of Engineering

Contents

Comparative Analysis of Hardware Performance Across ARM and x86 Architectures Using Computer Vision

Abstract—**The system proposed in this paper is a technology capable of matching a human face from a digital image or a video frame against a collection of known faces to find a match. The technology behind this is Facial Recognition, which is a contemporary security solution which automatically monitors, authenticates and identifies an individual's identity from these digital elements.**

This project will evaluate the performance between an x86 laptop with AMD Ryzen 7 5800H processor, a desktop with an Intel i7 13700k with a RTX 3080, and a ARM based architecture, the Raspberry Pi 4 a single-board computer. Both of these architectures will use software written programs in Python language and Software Development Kit integration. This project's program is then executed in each architecture respectively while python scripts and software are used to record the time it takes to verify within a video frame. This implementation is made possible with Computer Vision techniques using OpenCV Library version 4.10.0.84 with the Raspberry Pi 4 and interfacing with various tools such as a laptop webcam, database local storage device (SSD), standardized digital image formats such as JPEG or PNG for image storage, and use case verification confirming an individual's identity such as their name.

This project's evaluation of metrics will rely on active surveillance video and compare a database containing digital images to verify a person's identity. This project has met all objectives derived with optimized system performance and accuracy for the specific use case and environment. This project is considered successful and ready to be launched in the real system implementation.

I. Introduction

Biometrics scans various physical and behavioral characteristics to identify or verify an individual's identity. These biometric traits are used in various applications such as surveillance, law enforcement, healthcare and financing. Some of the most common biometric scans include fingerprint scanning, voice recognition, heart rate and rhythm, and facial recognition. Our subject area of research will focus on Facial recognition and the implementation of Computer vision techniques using OpenCV library version 4.10.0.84. Facial recognition is a way of identifying a face using this technology and biometrics. It uses video and/or picture elements from a source and goes against a collected database of known faces to find a match. Facial recognition systems are all around us in today's world, when we go grocery shopping, attending concerts, venues, and even our mobile phone we carry with us everyday in our pockets.

With the open source platform called OpenCV, machine learning has never been more accessible and can be done reliably. The Open Source Software OpenCV will collectively work with our compatible hardware component the Raspberry Pi 4, a Single Board Computer, to be able to detect human faces, compare the individual's facial signature to its database, and successfully identify you when present in front of the video camera. Additionally, our team will work with another hardware component, an x86 Laptop and follow the same procedure similarly with the Raspberry Pi 4 using the same Open Source Software and compare the two different architectures. In this report our team will address the benchmark algorithms between the x86 Laptop and the Raspberry Pi 4 with emphasis on 'the key metrics under examination include execution time, CPU usage, and memory utilization, with a focus on comprehending their impact on overall performance and overall cost analysis.

II. Architecture

A. ARM

To effectively evaluate the performance of the ARM architecture on the Raspberry Pi 4 in comparison to the x86 architecture in the AMD Ryzen 7 laptop, it is crucial to first understand the fundamental differences between Reduced Instruction Set Computing (RISC) and Complex Instruction Set Computing (CISC). In this section, we will focus on the ARM-based RISC architecture powering the Raspberry Pi 4, setting the stage for a detailed comparison with the CISC architecture of the AMD Ryzen 7 5800H and Nvidia RTX 3050ti laptop, which will be explored in the following section.

ARM (Advanced RISC Machines) is a widely adopted family of processors based on the principles of Reduced Instruction Set Computing (RISC). RISC architecture is designed to simplify the execution of instructions by using a small, optimized set of commands, with each instruction performing a single operation. This approach contrasts with Complex Instruction Set Computing (CISC) architectures, like x86, which use more intricate instructions that can carry out multiple operations simultaneously. The simplicity of RISC allows ARM processors to achieve greater efficiency and faster execution speeds, which are key advantages in energy-sensitive applications.

The ARM architecture's focus on minimizing the number of cycles per instruction contributes to its lower power consumption compared to other processor families, such as x86. This makes ARM-based processors, such as those found in the Raspberry Pi 4, highly suitable for embedded systems and mobile devices where power efficiency is critical. With its power-efficient design, ARM processors can perform complex tasks while consuming less energy, which is vital for systems like facial recognition, where real-time processing is required.

In addition to power and performance efficiency, ARM's modular design allows it to be easily integrated into various systems, such as embedded devices or single-board computers like the Raspberry Pi 4. This integration flexibility allows developers to implement facial recognition systems, utilizing the computational power of ARM processors to run machine learning models like OpenCV's Haar Cascade algorithms for real-time image processing. Overall, ARM's RISC architecture and the Raspberry Pi 4's optimization make it a competitive choice for applications requiring efficient performance, such as biometrics and computer vision.

B. x86

The x86 architecture is a family of instruction set architectures (ISA) originally developed by Intel and now used widely across various processors. Based on Complex Instruction Set Computing (CISC), CISC processors are designed to handle complex instructions that can perform multiple operations while taking several clock cycles to execute, which is the main difference between RISC and CISC processors. CISC processors, while capable of handling highly complex operations, typically require more power and generate more heat, often becoming limited by these factors.

The AMD Ryzen 7 5800H, an x86-based processor, exemplifies the power and versatility of the CISC design. It provides high clock speeds and supports multithreading, making it suitable for handling resource-intensive tasks. Additionally, CISC architecture includes a broad and sophisticated instruction set, enabling the processor to handle diverse workloads with fewer instructions in some scenarios, but with greater complexity per instruction. When paired with the Nvidia RTX 3050 Ti GPU, the system gains an edge in processing parallel tasks like those required for video and image processing. The GPU offloads these highly repetitive and intensive operations, such as applying OpenCV's Haar Cascade algorithms for facial recognition, allowing the CPU to focus on managing overall system performance.

One of x86's notable strengths is its extensive compatibility with a wide range of software and libraries, a benefit of its long history as a computing standard. This compatibility allows developers to utilize powerful tools for applications like facial recognition, ensuring efficient implementation and optimization. While CISC processors may require more clock cycles per instruction compared to RISC processors, their ability to execute complex operations makes them ideal for performance-heavy workloads. The combined capabilities of the AMD Ryzen 7 and Nvidia RTX 3050 Ti exemplify this, delivering a balance of computational power and flexibility that supports high-performance applications in biometrics and computer vision.

III. Hardware

A. Raspberry Pi 4

The Raspberry Pi 4 is a compact yet powerful single-board computer, making it ideal for diverse applications such as facial recognition. It features a quad-core ARM Cortex-A72 processor running at 1.5 GHz, providing sufficient computational power for demanding tasks while maintaining energy efficiency. The board includes up to 8GB of LPDDR4 SDRAM, dual-band 802.11ac Wi-Fi, Bluetooth 5.0, and gigabit Ethernet, ensuring robust connectivity. Additionally, it supports dual 4K monitor outputs via two micro-HDMI ports and comes with USB 3.0 ports for faster data transfers. This hardware design makes the Raspberry Pi 4 an excellent choice for integrating computer vision tasks such as those implemented with OpenCV. Its small size, modular design, and wide range of compatible accessories provide flexibility for both experimentation and deployment in real-world applications.

The MakerHawk USB Multimeter Voltage Tester is an essential tool for monitoring the performance of the Raspberry Pi 4. This versatile device measures voltage, current, temperature, and wattage with a calibrated digital output. With four operational buttons and five connection ports, it supports a range of USB types, including Type A, Type C, and Micro USB. Its two-way insertion detection and anti-interference capability ensure accurate measurements and stable performance. This multimeter simplifies testing and diagnosing issues with USB chargers, computers, and connected peripherals, enhancing the reliability of our setup.

The CanaKit Raspberry Pi 4 Starter Kit offers a complete setup to get the Raspberry Pi 4 up and running efficiently. It includes a high-gloss black case designed specifically for the Raspberry Pi 4, along with a set of three aluminum heat sinks and a low-noise fan to maintain optimal cooling during operation. The 3.5A USB-C power supply is UL-listed and incorporates a noise filter for improved stability. The kit also provides a 4K60P Micro HDMI to HDMI cable to support high-resolution dual-monitor displays. Known for its reliability and quality, the CanaKit Starter Kit ensures seamless integration of the Raspberry Pi 4 with all included components, making it a dependable solution for this project.

The Elecrow RC050 is a high-resolution capacitive touch LCD designed for use with the Raspberry Pi 4 and other mini PCs. It delivers an 800x480 resolution, ensuring clear and sharp visuals essential for monitoring and interacting with the system. This universal HDMI display includes a USB capacitive touch screen that supports five touch points without requiring additional drivers. Its intuitive OSD menu allows users to adjust settings, and the device features multiple ports for power, touch output, and HDMI connectivity. With its energy-saving backlight control and compatibility with a range of devices, the Elecrow RC050 enhances the usability of the Raspberry Pi 4 system, making it easier to manage and monitor the facial recognition application.

By combining the Raspberry Pi 4 with these high-quality accessories, this hardware setup ensures a robust, efficient, and reliable platform for implementing and evaluating facial recognition tasks.

B. Lenovo IdeaPad Gaming

The Lenovo IdeaPad Gaming Laptop offers a powerful platform for demanding tasks like facial recognition, due to its high-performance hardware. Powered by an AMD Ryzen 7 5800H processor with 8 cores and 16 threads, running at up to 4.4 GHz, the IdeaPad is capable of immense computation power using CISC processing. Paired with 16GB of DDR4 RAM, the laptop delivers fast memory performance and multitasking capabilities, ensuring smooth operation even under heavy workloads.

The laptop also features an NVIDIA GeForce RTX 3050 with 4GB of VRAM, providing powerful graphics processing power. This GPU accelerates tasks such as real-time image processing and machine learning, making it a highly suitable choice for running OpenCV's Haar Cascade algorithms. Additionally, the Lenovo IdeaPad comes with a 15.6-inch full HD display, ensuring notable visuals for monitoring and interacting with the system. With this combination of high-end CPU and GPU, the Lenovo

IdeaPad Gaming Laptop is a powerful and efficient tool for real-time facial recognition and other computer vision tasks. The laptop also includes a built-in 720p HD webcam, providing essential input for facial recognition tasks which will be used in the experiment to collect data. By leveraging this hardware setup, the laptop is well-equipped to handle resource-intensive applications, making it ideal for tasks like image processing, machine learning, and other computationally intensive activities associated with facial recognition and computer vision.

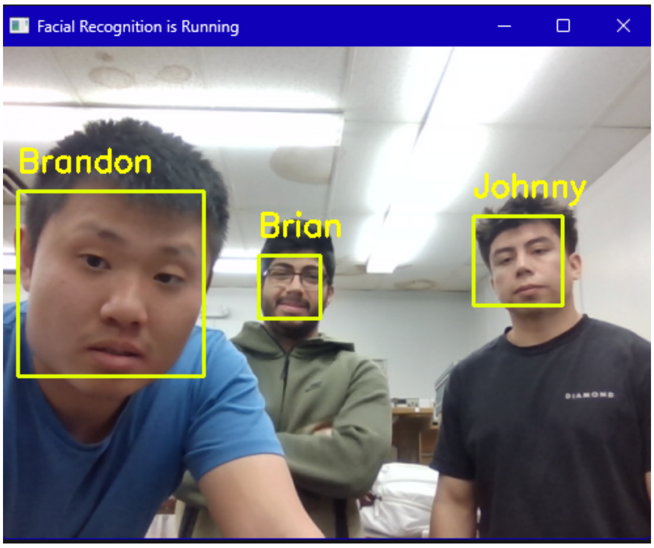
C. Desktop ((Intel Core i7 13700k, 32GB ram, RTX 3080 10GB VRAM)

The desktop configuration featuring the Intel Core i7-13700K processor is designed for high-performance computing. With 16 cores (8 performance cores and 8 efficiency cores) and 24 threads, this CPU offers exceptional processing power, capable of handling demanding applications like facial recognition and machine learning at high speeds. Its base clock speed of 3.4 GHz can be overclocked up to 5.4 GHz, providing fast, responsive performance for even the most complex workloads. Paired with 32GB of high-speed DDR4 RAM, the system ensures smooth multitasking and efficient data handling, allowing the desktop to manage large datasets and parallel processing without slowing down.

The inclusion of the NVIDIA GeForce RTX 3080 graphics card with 10GB of VRAM makes this system ideal for graphics-intensive tasks. The RTX 3080 is equipped with NVIDIA's Ampere architecture, offering enhanced performance for real-time image processing, 3D rendering, and running computationally intensive algorithms such as OpenCV Haar Cascade for facial recognition and convolutional neural networks for descriptors. Containing such a powerful GPU, the system can accelerate machine learning processes and deliver high-quality graphics and visuals, ensuring the applications of computer vision are most optimal.

IV. Algorithms

The algorithms that are chosen from the openCV library are Haar Cascade, Histogram of Oriented Gradients or HOG, and Convolutional Neural Networks or CNN written in python. The project utilizes two algorithms where the foundation of the system begins with the use of Haar Cascade which serves as the initial face detection method with pre-trained models, it works by dividing an image into small cells and calculate the gradient direction and magnitude for each pixel within the cells and then utilizes either the HOG or CNN algorithm [1]. The latter two algorithms work with Haar Cascade, and as soon the initial detection is done the second algorithm will be enabled for refined detections. For the x86 systems they ran python version 3.10.0 a stable version where it supports Nvidia's deep learning tools and the Raspberry Pi 4 using the Buster OS supporting python version 3.7. If Haar Cascade is deployed with HOG it is used to detect faces by focusing on the edge and gradient patterns within the images, it works by rescaling and converting the images to vectors or array lengths [2]. The rescaled image helps calculate the different sizes of arrays and vectors to compute the gradients across the image to see the changes in the pixel values to get important features. This allows HOG to describe the overall shape of objects while being computationally efficient.



When the project uses CNN with Haar Cascade it offers a more powerful and accurate approach for face detection and recognition [2]. This technique automatically learns from hierarchical features from the image during training eliminating the need for handcrafting the extraction. It processes the images through a series of convolutional, pooling, and connected layers to detect and extract features like eyes, nose, and mouth.

Now with these algorithms they can run on both x86 and ARM processors. To benchmark and compare the two architecture, the x86 system is a Ryzen 7 5800H an eight core processor with a RTX 3050 ti laptop to the ARM based processor the Broadcom BCM2711 a four core processor, a second x86 system was also included with a Intel i7 13700k with a RTX 3080 desktop and set the processing algorithm for both to the HOG model to see the time requirement it takes for the processor to train the image and begin the facial recognition procedure. To measure the power usage for the x86 systems a software was used instead of a physical power meter because the meter would measure the total system power instead of the only the processor itself which is what is being compared between the the two platforms so HWMonitor software was used to only track the processors.

V. Performance Evaluation

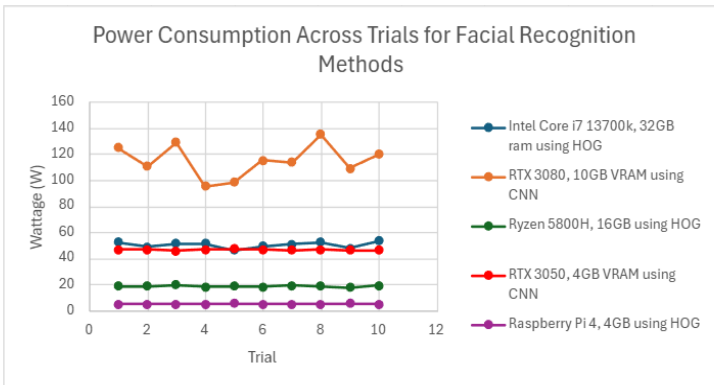


Figure 1: Power Consumption Across Trials for Facial Recognition Methods

In figure one for the x86 architecture the desktop system with the i7 13700k with the RTX 3080 consumes the most power, followed by the laptop system a Ryzen 7 5800H with the RTX 3050, and finally the most power efficient RPi 4. The desktop system has the highest total system power, but using the software HWMonitor the CPU and the GPU wattage can be isolated so the comparison between the other processors can be consistent. Using the CNN model the GPU consumes the most amount of power at an average of 118.43 watts, this algorithm has the highest computational process that utilizes the GPU that has 8704 cuda cores at 1800 MHz. These cores run in parallel and there are multiple times more cores than regular CPU hence the high power draw. With the addition of video RAM or VRAM with a high bandwidth of 760 GB/s it can process high amounts of data. The GPU is followed by the i7 13700k the desktop CPU and the RTX 3050 the laptop GPU. The desktop CPU averages 50.18 watts this processor has 16 cores and 24 threads where the cores are split with 8 performance cores and 8 efficiency cores at 5.4 GHz and 4.2 GHz respectively, the HOG model was used to train the local data and run the facial recognition model. The HOG model was used instead of the CNN model because HOG is designed for simple vector based calculations instead of the complex dot products and matrix calculations. The CNN model may be used with the CPU, but the complex computation makes it take significantly longer to train and run the facial recognition model. Now the next system, the laptop RTX 3050 consumes 46.72 watts. This GPU also used the CNN model as well, but it is the lower end part for the 3000 series Nvidia GPU. The 3050 has 2048 cuda cores at 1545 MHz, but is capable of running the CNN model effectively at a slightly lower power level than the 13700k. The next part is the Ryzen 5800H an 8 core 16 thread at 3.2 GHz a laptop version of the 5800 series of CPU, it is a more power efficient version of it for battery powered applications, and finally the RPi 4 the most power efficient system with the BCM2712 quad core processor at 1.8 GHz ARM based processor.

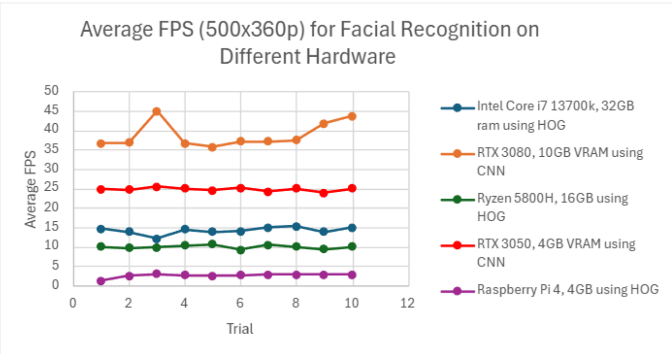


Figure 2: Average FPS (500x360p) for Facial Recognition on Different Hardware

In figure two this data set is plotting the average frames per seconds or fps when the facial recognition model is running. From the data the most fastest and power hungry part is the RTX 3080 averaging at 38.71 fps, using the CNN model the graphic card has thousands of cores that can run in parallel as we covered before and with the addition with VRAM and a high memory bandwidth it can detect and describe faces more effectively and accurately. The 3080 is followed up by the entry level graphic card the 3050 where it averages at 24.73 fps, even though it has a fractional amount of cuda cores, VRAM, and memory bandwidth it is still effective enough to run the CNN model at a slightly lower power level as the high end i7 13700k CPU where it only averages at 14.11 fps with the HOG model. Even though it has a significantly higher clock rate at 5.4 GHz for the performance core and 4.2 GHz for the efficiency cores it is unable to run the CNN model effectively and won't be faster than the graphics cards even though it is running the less complex image descriptor model. The CPU is followed by the Ryzen 7 5800H averaging at 9.91 fps and then finally the ARM based processor BCM2712 for the RPi 4 averaging at 2.53 both running under the HOG model. Even though the RPi 4 has the lowest performance metrics it is impressive that it can still run the facial detection model with such efficiency following the trend for RISC based architecture.

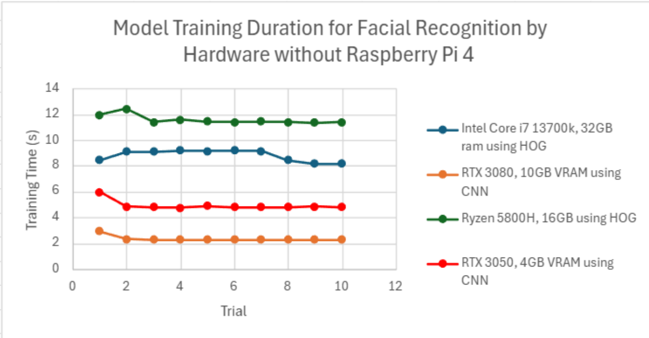


Figure 3: Model Training Duration for Facial Recognition by Hardware without Raspberry Pi 4

In this figure it is the training portion of the experiment, the project is built locally so the images that are provided are the group members' photos so the facial recognition models can put descriptions on who is on the frame. From the data on the RTX 3080 using the CNN model we see that it is the fastest in training the data set that we have provided. With 87 images the 3080 is able to process the images in an average of 2.82 seconds. This is followed by the next graphic card, the 3050 that trained the data set in an average of 4.95 seconds. Now with the CPU they can train with the CNN model but as said before the training would be not feasible because it would take over an hour to train so the HOG model is used instead which is optimized for CPUs. Without the RPi 4 we wanted to first test all the x86 systems first because the difference between the systems is so great other data cannot be shown on the plot. But with the other CPU the 13700k comes in the third fastest at an average of 8.79 seconds and the Ryzen 7 5800H at 11.56 seconds. From the data we can see how beneficial it is to have thousands of cores from the GPUs that gives it the capabilities to train quickly on the CNN model. With the smaller core counts of CPUs they have to run on the HOG model although it is optimized for CPUs they are still not fast enough to complete the training compared to the GPU counterparts.

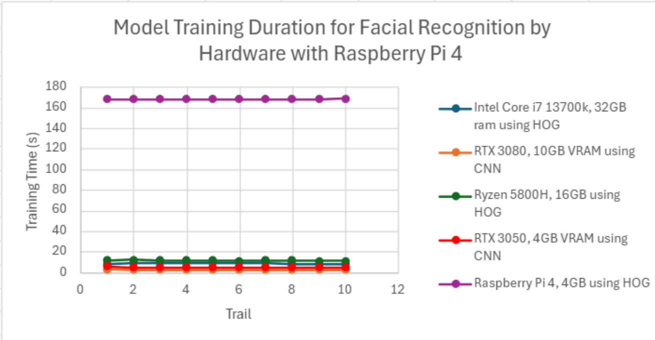


Figure 4: Model Training Duration for Facial Recognition Hardware with Raspberry Pi 4

Now in figure 4 the RPi 4 is included in the data, this had to be separated because the RPi 4 is so slow at training the data it makes the other plots have difficulty showing up on the plot. But with the ARM based BCM2712 quad core processor it took an average of 167.95 seconds which is significantly greater than the x86 systems, however the RPi 4 has is a power efficient system that still has the capabilities to run the model, it is still effective enough to train the models and run the facial recognition models. This system is effective for power efficient consumer grade work not for large power hungry enterprise systems.

All the systems are capable of running the training algorithm and the facial detection algorithm, but it depends on the use case. If it is for an enterprise where a powerful x86 system is needed to run facial recognition software then that type of system is needed to run that algorithm to have accurate detection results. It can be used for employee tracking or large data bases that companies need to handle. If power is a constraint and lets say that it is in a battery powered application then the RISC based ARM architecture is best suited for the situation because it is a low power system that can run these algorithms as well but at a significantly lower power.

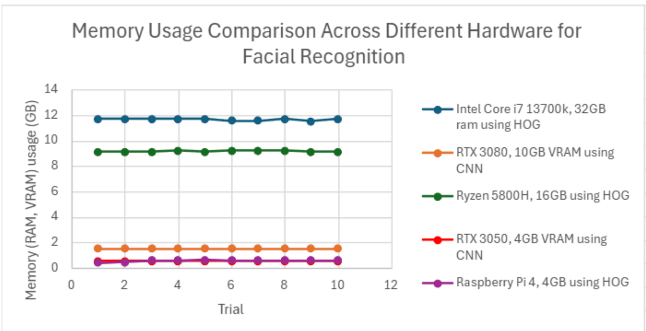


Figure 5: Memory Usage Comparison Across Different Hardware for Facial Recognition

In figure 5 we are comparing the DRAM usage for the CPUs running the HOG model and the VRAM usage for the GPUs on the CNN model. In the plots we see that the i7 13700k consumes approximately 11.69 GB of RAM while the 5800H consumes 9.15 GB of ram, and the ARM based system for the RPi 4 consumes roughly 0.5 GB of RAM. From this analysis the RPi 4 consumes significantly less memory compared to the x86 counterparts, this is due to the efficiency of the ARM based processors for lightweight operations, with the addition with the HOG model that is optimized for CPUs it is energy efficient for compact processing. On the other hand the x86 systems are much higher performance but at the cost of higher memory and power consumption, since the HOG model is less computationally demanding it requires higher memory allocation to handle the facial recognition process due to the reliance of general purpose processing. The GPUs run on the CNN model which is computationally higher, however the 3080 only consumed approximately 1.5 GB of VRAM and the 3050 only consumed 0.5. This is due to the resolution scaling of the project, in the python script a variable is used to control the resolution, the value of 500 was used because it is any higher the RPi 4 would struggle to run the HOG models. If the variable value were to increase to 1000 it would consume additional VRAM memory but it would have additional power draw for the GPUs because of the intense computational processing.

VI. Conclusion

The comparative analysis of ARM and x86 architectures in the context of computer vision and facial recognition demonstrates both the versatility and specific capabilities of these different computing platforms. Throughout our evaluation, we observed significant differences in terms of execution time, power efficiency, and computational effectiveness across the ARM-based Raspberry Pi 4, the x86-based Lenovo IdeaPad with AMD Ryzen 7, and the desktop system with Intel i7 and RTX 3080.

The ARM architecture, embodied by the Raspberry Pi 4, showed strengths in power efficiency and cost-effectiveness, making it well-suited for environments where energy consumption is a constraint, such as embedded or portable systems. However, it was clear that the ARM system struggled to achieve comparable computational performance to the x86 platforms, particularly in tasks involving heavy real-time processing like model training for facial recognition.

On the other hand, the x86 platforms—particularly the desktop with Intel i7 and RTX 3080—excelled in terms of raw performance, with higher frame rates, quicker training times, and more effective handling of high-resolution image data. The combination of a high-performance CPU and GPU in this system enabled efficient parallel processing and rapid model execution, which are crucial for high-demand tasks. The AMD Ryzen 7 system also provided respectable performance, effectively balancing computational power with energy usage, making it a practical middle-ground solution for many applications.

Ultimately, the choice between these architectures depends on the specific requirements of the intended application. For use cases requiring mobility, lower power consumption, and sufficient real-time processing capabilities, ARM is a viable choice. In contrast, for applications demanding high processing speed, accuracy, and the ability to handle intensive workloads, x86-based systems are more appropriate. The analysis emphasizes the trade-offs between efficiency and performance, providing insights into selecting the right hardware based on context, budget, and functional requirements.

VII. References

[1] S. Saha, "Histogram of Oriented Gradients (HOG)," LearnOpenCV, Oct. 10, 2024. [Online]. Available: https://learnopencv.com/histogram-of-oriented-gradients/. [Accessed: Nov. 25, 2024].

[2] C. Rahmad, R. A. Asmara, D. R. H. Putra, I. Dharma, H. Darmono, and I. Muhiqqin, "Comparison of Viola-Jones Haar Cascade Classifier and Histogram of Oriented Gradients (HOG) for Face Detection," IOP Conference Series: Materials Science and Engineering, vol. 732, p. 012038, 2020, doi: 10.1088/1757-899X/732/1/012038.

[3] R. A. Asmara, M. Ridwan, and G. Budiprasetyo, "Haar Cascade and Convolutional Neural Network Face Detection in Client-Side for Cloud Computing Face Recognition," 2021 International Conference on Electrical and Information Technology (IEIT), 2021, pp. 1–6, doi: 10.1109/IEIT53149.2021.9587388.