GPMS: A Generative Projection Mapping System

Figure i *Projection Concept* [1]

University of Central Florida
Department of Electrical and Computer Engineering

Senior Design I
Dr. Chan

Group B
Victoria Moreno - Computer Engineering
Declan Carter - Computer Engineering
Francisco Soriano - Computer Engineering

Committee
Reza Abdolvand, Saleem Sawahneh, George Atia

Table of Content

List of Figures

List of Tables

Chapter 1: Executive Summary

The project introduced by the authors is a Generative Projection Mapping System (*GPMS*), an intriguing technology aimed at delivering projection mapping for everyday use. The authors desire to make immersive storytelling and entertainment accessible to hobbyists and entertainers alike. The GPMS initiative seeks to bridge the gap between high-end, professional projection mapping displays utilized by entertainment giants and the consumer market. The motivation is fueled by the potential to enhance personal and communal experiences through the innovative use of generative AI and computer vision algorithms.

GPMS is designed as a portable system capable of transforming any surface into a dynamic visual storytelling canvas. By accepting user inputs, GPMS leverages generative AI to produce images that are then projected onto surfaces, aligning with their unique features. The device is engineered to recognize the structure's vertices, edges, and designs, ensuring that generated images perfectly complement the physical space. GPMS combines a projector with a responsive touchscreen monitor for easy use and advanced image processing algorithms for seamless image integration.

The primary goal of GPMS is to offer a user-friendly, customizable projection experience that allows for a high degree of personalization in how images are matched and displayed against targeted structures. Key objectives include both ensuring the projector accurately displays images based on user prompts and allowing for precise alignment of generated images with the structure's specific features.

The GPMS system has broad applications that promise to revolutionize how stories are told and how experiences are shared among the general public. Furthermore, this system will make projection mapping an accessible tool for creativity. The project's ambition extends beyond its current scope, with plans to collaborate with local theme parks and potentially establish a startup, significantly impacting the entertainment industry and consumer technology.

GPMS aims to set a new and unique standard in immersive visual storytelling. This system represents a significant step towards making advanced projection mapping accessible, offering a blend of creativity, innovation, and personalization that has the potential to enrich everyday experiences.

Chapter 2: Project Description
2.1 Motivation and Background

The authors' interest in projection mapping stems from their passion for immersive storytelling and entertainment. However, due to the artistry, time, and technical nuance often required to tailor-make projection mapping shows, such displays are typically beyond the grasp of everyday hobbyists or entertainers. While theme park giants like Disney and Universal continue to improve large, multi-projector systems to render grand scenes onto their existing architecture, more needs to be done to put this technology into everyday consumers' hands. Individuals already look to purchase Sky Projectors to throw the night sky onto their ceiling, Christmas lights to celebrate the holidays, or disposable paper products to decorate their walls for festive occasions. However, by combining state-of-the-art computer vision algorithms and generative AI models, the authors look to introduce a new and unique product to allow people to share the joy of any occasion.

The authors propose a portable generative projection mapping system, *GPMS*, that can map generated images onto any unique surface that will create a themed and immersive environment. This system will use generative AI to create themes that precisely fit user input. This is done by GPMS aligning the major image features with the existing vertices on the surface. This process will be scalable using a central and distinct server to generate these images. As a result, the system can be employed relatively cheaply. The generative AI aspect of GPMS allows users to let their audiences experience immersive storytelling.

The authors take inspiration from the historical presence of Projection Mapping in local entertainment venues. In 2011, Disney World, located in Orlando, Florida, began projecting images onto Cinderella's castle with the help of 3D scanning and projection mapping. The emerging projections ran from 2011 to 2012 and even displayed visitors' photos onto Cinderella's castle throughout select shows. In light of the increase in the versatility and stability of projection mapping, Disney World has continued incorporating more creative and new projection shows for years to come. [2]

Additionally, Universal Orlando has transformed its projection mapping techniques through two major breakthroughs within the last four years. The first breakthrough was the development of high-resolution projectors capable of beaming bright and crystal-clear images over long ranges. The second breakthrough was projection mapping to ensure that projected images were not distorted when beamed onto uneven structures such as the Hogwarts Castle, one of Universal's more notable attractions. [3]

As entertainment venues increase their use of projection mapping, a unique business opportunity arises for the authors. More specifically, after the final demonstration, the authors plan to propose GPMS with a buy option to local entertainment venues in Orlando, FL, with the condition that the projector must be returned unharmed. Additionally, the authors foster a hope to create their own startup from scratch shortly after the demo.

Projection mapping, being relatively recent, has been refined to project many images that can wrap themselves around various surface types for the images to look like they are part of the surfaces themselves. As projection mapping advancements continue to grow and expand, GPMS aims to directly engage and provide storytellers in all environments with a fresh way of delivering messages through the use of Generative AI.

2.2 Related Work

2.2.1 Dynamic Projection Mapping Research

A projection mapping solution presented by Addison Sandvik of the University of California Polytech State University outlines a similar approach to GPMS. This project demonstrates a proof-of-concept for a flexible projection mapping technique, allowing a surface to be placed anywhere within the projector's field of view for accurate image projection. The technique employs a surface equipped with infrared lights, and a camera outfitted with an infrared filter to identify the location of the surface. Then, this technique utilizes a projective transformation to convert points from the camera's perspective into corresponding points in the projector's output. [4]

Sandvik's work demonstrates a few basic implementations of computer vision aspects, such as lens distortion correction using a pre-calculated matrix to flatten the rounded image and image calibration/alignment by transforming the corner points of the projected image to where the camera is able to view them. Both of these implementations are essential aspects of the computer vision stack necessary for GPMS. Additionally, in a few key ways, the authors' implementation simplifies a few critical aspects as their projection surface is *static*, unlike the dynamic mapping proposed by Sandvik. So, some pieces of Sandvik's project can be well adapted to suit the needs of any calibration issues that GPMS may experience.

2.2.2 Industry Products: Luma Box

Many products in the industry do projection mapping, and some are even AI-generated. The best existing example of a product like this is Luma Box. Luma Box introduces an innovative approach to projection mapping through its product, Luma Map. This tool is designed to facilitate the creation of visually engaging projections that can be tailored to fit the unique contours of any surface, such as that of a house. The versatility of Luma Map lies in its ability to enhance physical features with light and shadows while integrating AI-generated themes.

Moreover, Luma Map presents a relevant case study for a project that employs AI to generate images for projection onto buildings, thereby augmenting their appearance or introducing thematic elements. It exemplifies how technology can merge seamlessly with generative AI to redefine spaces, offering insights into the potential applications and benefits of similar AI-driven projects in projection mapping. [5]

2.2.3 Stable Diffusion

Stable Diffusion is a state-of-the-art deep learning model developed by Stability AI that has revolutionized the field of text-to-image generation. It is a latent diffusion model trained on a vast dataset of text-image pairs, enabling it to generate highly realistic and diverse images from textual descriptions. [6]

At its core, Stable Diffusion is based on a type of generative model called a diffusion model. Diffusion models work by learning to reverse a gradual noising process applied to training images. During the training phase, the model is shown images that are progressively degraded with noise, and it learns to recover the original clean images from the noisy versions. [^]

The key innovation of Stable Diffusion lies in its ability to combine this diffusion process with a powerful text encoder, such as the Transformer architecture, which allows it to understand and translate textual descriptions into meaningful visual representations. The model learns to associate specific words and phrases with visual concepts, enabling it to generate images that accurately reflect the content and style described in the text prompt. [6]

Stable Diffusion employs a technique called latent diffusion, which operates in a compressed latent space rather than directly on pixel values. This approach enables the model to capture high-level semantic information and generate images with improved quality and efficiency compared to previous text-to-image models. [6]

The architecture of Stable Diffusion consists of an encoder, a series of diffusion steps, and a decoder. The encoder maps the input text into a latent representation, which is then iteratively refined through the diffusion process. At each step, the model predicts the noise that needs to be removed to reconstruct the original image. The decoder then maps the final latent representation back into the pixel space, generating the output image. [6]

One of the key advantages of Stable Diffusion is its ability to generate diverse and creative images from a given text prompt. By sampling from the learned distribution of the model, it can produce multiple plausible visual interpretations of the same textual description. This allows users to explore a wide range of creative possibilities and generate unique and compelling images. [6]

Stable Diffusion has found applications in various domains, including art, design, gaming, and visual storytelling. It has the potential to assist artists, designers, and content creators in generating novel and inspiring visual content, as well as enabling new forms of creative expression and exploration.

2.2.4 Stable Diffusion with ControlNet

While Stable Diffusion excels at generating images based on textual prompts, it can be challenging to accurately convey specific visual elements or spatial arrangements through words alone. To address this limitation and provide users with more control over the generated images, a group of researchers from Stanford University introduced ControlNet in 2023.

ControlNet is “a neural network architecture that enables the incorporation of spatially localized input conditions into a pre-trained text-to-image diffusion model, such as Stable Diffusion, through efficient fine-tuning (Zhang et al., 2023).” By allowing users to provide an additional input image alongside the text prompt, ControlNet enhances the user's ability to guide the image generation process and achieve more precise results. For example, if a user wants to generate an image of a superhero in a specific pose, providing an input image of the desired pose would be more effective than attempting to describe the pose in words. ControlNet leverages this visual information to constrain and direct the image generation process, resulting in outputs that more closely align with the user's intentions. [7]

A key aspect of ControlNet is its use of a hypernetwork architecture. Instead of retraining the entire Stable Diffusion model from scratch, ControlNet’s makers were adamant in reusing the model. ControlNet acts as a hypernetwork, a neural network that generates weights for another network. [8] This approach allows for efficient adaptation and specialization while preserving the knowledge and capabilities of the original model.

The diagram below, taken from the researchers' paper, illustrates the architecture of ControlNet and its integration with Stable Diffusion:

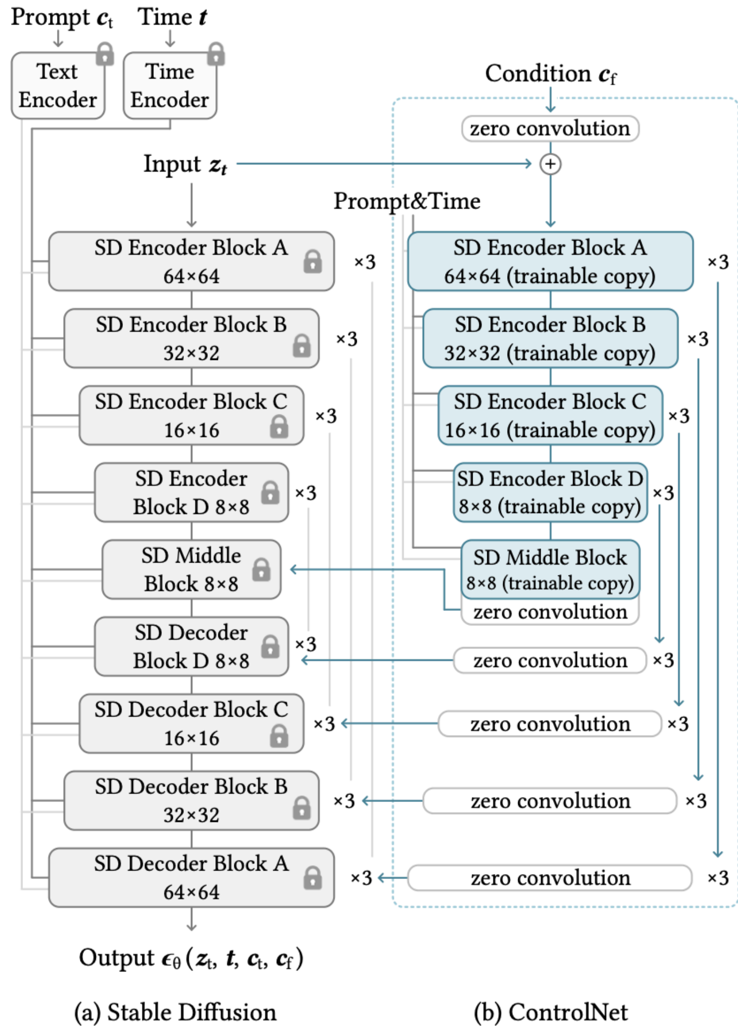


Figure 2.2.4 *Stable Diffusion with ControlNet*

In this architecture, ControlNet consists of a trainable copy of the encoders and the middle block from the Stable Diffusion model. The Stable Diffusion model itself remains locked, ensuring that its pre-trained weights are not altered during the fine-tuning process. ControlNet utilizes zero convolution layers in its initial stages to mitigate the impact of harmful noise during training. The weights generated by ControlNet's layers are then passed to the decoders of the Stable Diffusion model. The middle block of ControlNet also contributes weights to the corresponding middle block of Stable Diffusion, which are summed together before propagating to the subsequent decoder blocks. This iterative process of combining weights from ControlNet and Stable Diffusion continues through the decoder blocks until the final output image is generated.[7]

By selectively injecting learned weights into specific parts of the Stable Diffusion model, ControlNet enables more precise control over the image generation process while leveraging the pre-trained model's knowledge and capabilities. The results presented in the researchers' report demonstrate the effectiveness of ControlNet in enhancing the controllability and flexibility of Stable Diffusion. By preserving the essential features of the input structure and incorporating them into the generated image, ControlNet opens up exciting possibilities for user-guided image generation across various domains, including art, design, and creative applications. The combination represents a significant advancement in the field of generative AI, empowering users with greater control and expressiveness in creating visually coherent and semantically meaningful images.

2.3 Project Overview

2.3.1 Basic Goals

The authors’ project is centered around creating a system, GPMS, with the primary goal of constructing an immersive storytelling environment. The overarching objective is to dynamically generate and project images onto a designated structure in alignment with its unique features. These structures' features include but are not limited to pillars, windows, bricks, or even just a drawing. Therefore, the camera utilized within the system must be of sufficient quality to best capture the features. GPMS will adeptly detect vertices, edges, and designs on the given target structure by intelligently leveraging sophisticated computer vision techniques.

GPMS integrates a responsive touchscreen monitor to facilitate seamless user input and display the respective generated content. Users can input text/prompts, which will directly initiate the generation of images that will then be intelligently projected onto the unique structure/surface. The authors’ cutting-edge system harnesses the power of generative AI to interpret user inputs, along with capabilities to craft images that align seamlessly with the unique structure's features.

For example, during the winter, one may ask for a Christmas design dedicated for a house. As a result, the generated image could include candy canes for the front pillars and snowflakes on the featureless areas. Alternatively, if the surface was a drawing of a dog, the generated image could include that dog with a Santa hat on its head.

GPMS must have a high-powered projector for optimal visibility. In order to calibrate the image the user must manually select the edges of the projected display each time it is used on a new structure. User calibration is achieved through the touch screen by moving the image's corners to the proper areas on the structure. The user will know if the calibration is set correctly because the image will be projected on the structure as they are calibrating.

2.3.2 Advanced Goal

In the author's pursuit of elevating user customization, the overarching goal is to increase the control users have over image representation, aiming for a significant level of precision. Their specific objective is the implementation of sensitivity and threshold controls to achieve this goal. These controls will empower users, allowing them to intricately adjust the outline of the structure according to their unique preferences. For example, if the structure was a brick wall, but the user did not want their projected image to be based on each brick's outline, they can turn the sensitivity down. However, if the structure is a drawing of someone's face, the user can increase the sensitivity, allowing for the inclusion of each specific feature on the face.

This transformative experience is facilitated through a user-friendly interface, where individuals can precisely determine the desired similarity between the generated image and the structure. The innovative approach proposed by the authors ensures a seamless and personalized customization experience, setting a new user engagement and satisfaction standard.

2.3.3 Stretch Goals

As part of the authors' stretch goals, they aim to allow users to highlight specific structural features. Their primary objective is to design an intuitive interface that allows users to effortlessly select and emphasize particular elements, ensuring a personalized visual experience. Users can achieve this by choosing features on the picture of the structure through the monitor. The image generated will closely align with the selected features.

Their second goal is to streamline the user experience. The objective is to integrate an automatic calibration process in the device, adapting to the features of walls or selected structures with minimal user input. Autocalibration ensures optimal image display without unnecessary complexity. Given these enhancements, all the user must do is set up the device, and the image will adequately align with all the features of the structure. Below is a summarization of the specific types and descriptions of each of the authors' objectives.

- **Basic Objectives**
    - The projector seamlessly displays an image corresponding to the user input.
    - The generated image intelligently aligns with distinct features on the targeted structure.
- **Advanced Objective**
    - Implementation of an input method to adjust sensitivity and threshold for image outlines.
- **Stretch Objectives**
    - Develop a user-friendly frontend that empowers users to highlight

features on the structure for emphasis.
  - Automatic calibration of the device to the specific features present on the wall.

2.4 Requirements Specification

2.4.1 Target Project Specifications

The target specifications, which are listed in table format below, are directly displayed in the House of Quality (HOQ) below. The specifications in the table below were derived with the main motivation of assembling GPMS as efficiently as possible.

Category	**Target**
Size	12 inches x 12 inches x 5 inches
Weight	8 pounds
Power Transmission Efficiency	~1 hour
Display Accuracy	100% accuracy
Response Time	< 45 seconds
Visual Quality	1920 x 1080 pixels
Functionality	< 10% failure rate
Projection Range	> 6 feet
Calibration Autonomy	100% Fully Autonomous*
Recover Time	< 10 seconds

Table 2.4.1 *Target Project Specifications*

*Note: Complete autonomy is an advanced goal.

2.4.2 Response Time

Achieving a response time of less than 45 seconds within GPMS is paramount for ensuring an immersive and dynamic storytelling experience. The system's ability to swiftly generate and project images onto designated structures hinges on its responsiveness to user inputs. With a rapid response time, users can seamlessly interact with the touchscreen monitor by inputting text or prompts to trigger the generation of images tailored to the unique features of the target structure. Whether it's pillars, windows, or even intricate drawings, GPMS must swiftly interpret user commands and deploy generative AI algorithms to craft visually stunning projections in real-time.

The importance of a swift response time extends beyond just user convenience. In fact, it directly impacts the system's effectiveness in delivering tailored content that aligns seamlessly with the structure's features. By harnessing sophisticated computer vision techniques, GPMS can intelligently detect vertices, edges, and designs, ensuring that generated images precisely fit the contours of the target surface. This precision is extremely essential for creating an immersive storytelling environment where each and every projection feels integrated and cohesive, captivating all audiences with its seamless blend of digital and physical elements.

Furthermore, a rapid response time within GPMS is crucial for facilitating user calibration of the projected images onto the unique structure. With the ability to manually adjust the image's position in real-time via the touchscreen interface, users can ensure optimal visibility and alignment. Whether it's adjusting the corners of the projected image to match the contours of a building facade or fine-tuning the placement of elements within a drawing, a swift response time empowers users to effortlessly customize their projection mappings with precision and accuracy. Ultimately, by prioritizing a response time of less than 45 seconds, GPMS enhances user engagement, fosters creativity, and delivers an unparalleled immersive storytelling experience.

2.4.3 Visual Quality

Ensuring a visual quality of 1920 x 1080 pixels within GPMS is paramount for delivering immersive and high-fidelity projections that captivate audiences and convey content with clarity and precision. With a resolution of 1920 x 1080 pixels, GPMS can render various images and graphics with exceptional detail, allowing for the faithful reproduction of intricate designs, vibrant colors, and subtle textures. Whether it's projecting dynamic animations, lifelike visuals, or informative graphics, the system's high-resolution output ensures that every pixel contributes to a visually stunning experience that leaves a lasting impression on all viewers.

Moreover, a visual quality of 1920 x 1080 pixels enables GPMS to accommodate a wide range of content types and formats, from high-definition videos to detailed graphics and text. This versatility is essential for catering to diverse creative visions and content requirements, whether the system is used for artistic installations, corporate presentations, or educational purposes in University settings. By delivering crisp and clear visuals with a resolution that meets or exceeds standard high-definition specifications, the GPMS ensures that content creators have the flexibility and freedom to unleash their creativity and realize their vision with uncompromising visual fidelity.

Furthermore, a resolution of 1920 x 1080 pixels enhances audience engagement and immersion by creating a seamless and immersive viewing experience that transports viewers into the heart of the content. Whether it's exploring virtual worlds, experiencing interactive narratives, or witnessing breathtaking visuals, audiences can fully immerse themselves in the projected content without distraction or loss of detail. This heightened level of immersion not only enhances the impact of the storytelling experience but also fosters a deeper connection between viewers and the content, leaving a lasting impression that resonates long after the projection has ended. By prioritizing visual quality of 1920 x 1080 pixels, the GPMS elevates the art of projection mapping to new heights, setting the stage for unforgettable visual experiences that push the boundaries of creativity and technology.

2.4.4 Projection Range

Ensuring a projection range of greater than 6 feet within GPMS is fundamental to accommodating diverse spatial configurations and delivering immersive visual experiences across various environments. With a robust projection range, the system can effectively cover larger surfaces, allowing for greater flexibility in venue selection and setup. Whether it's projecting onto expansive building facades, wide interior spaces, or outdoor landscapes, a generous projection range ensures that the GPMS can adapt to different spatial constraints while maintaining optimal image quality and clarity.

Moreover, a projection range exceeding 6 feet enhances accessibility and audience engagement by enabling projections to be visible from a decent distance. This is particularly crucial for outdoor installations, large-scale events, holiday events, gatherings, or venues with expansive seating arrangements. By reaching farther distances, GPMS expands its reach to a broader audience, ensuring that everyone within the viewing area can fully immerse themselves in the projected content. Whether it's a captivating visual narrative, interactive art display, or informative presentation, an extended projection range maximizes the system's impact and ensures that the intended message or experience reaches its audience effectively.

Furthermore, a robust projection range within GPMS opens up possibilities for creative expression and innovative storytelling techniques. Artists, designers, and content creators can leverage the system's extended reach to experiment with larger-than-life projections, immersive environments, and interactive installations that transcend traditional boundaries. From immersive art installations that transform entire cityscapes to interactive experiences that invite audience participation from afar, a projection range exceeding 6 feet empowers creators to push the boundaries of storytelling and engage audiences in unforgettable ways. By prioritizing a generous projection range, GPMS becomes a versatile tool for artistic expression, entertainment, education, and cultural enrichment, enriching the lives of audiences worldwide.

2.4.5 Recover Time

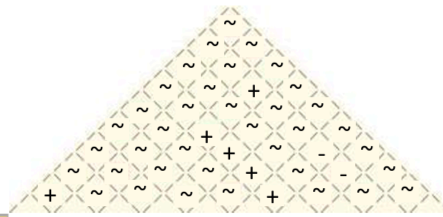
Maintaining a recovery time of less than 10 seconds within GPMS is essential for ensuring uninterrupted operation and seamless user experiences. In dynamic environments where technical glitches or interruptions can occur, a swift recovery time is paramount for minimizing disruptions and maintaining the continuity of the storytelling experience. Whether it's a momentary power outage, software glitch, or hardware malfunction, the ability of the GPMS to swiftly recover within 10 seconds ensures that users can quickly resume their interactions with the system without significant downtime or frustration.

Furthermore, a rapid recovery time is crucial for preserving audience immersion and engagement during live events or performances. In scenarios where the GPMS is used for real-time projection mapping during concerts, theatrical productions, or interactive installations, any downtime can take away from the overall experience and actually disrupt the audience's engagement momentum. By minimizing recovery time to less than 10 seconds, the system can seamlessly adapt to unforeseen challenges, ensuring that the show goes on without missing a beat. This reliability enhances the credibility of the GPMS as a professional-grade tool for immersive storytelling and live entertainment applications.

Moreover, a quick recovery time within the GPMS bolsters confidence among users and event organizers, fostering trust in the system's reliability and performance. Whether it's deployed for corporate presentations, experiential marketing campaigns, or cultural events, the ability of the GPMS to swiftly recover from technical hiccups instills peace of mind and reassures stakeholders that their investment in the technology is well-founded. By prioritizing a recovery time of less than 10 seconds, the GPMS demonstrates its resilience and readiness to overcome challenges, presenting itself as a dependable solution for delivering impactful visual experiences in any setting.

2.4.6 House of Quality

The House of Quality below depicts the Target Specifications listed in Table 2.4.1. In addition, Engineering Requirements, Marketing Requirements and their respective direction of improvements and relationship with each other are listed. The correlations are also listed in the relationship matrix/roof of the HOQ.



House Of Quality

	Marketing Requirements	Direction	1 Size	2 Weight	3 Power Transmission Efficiency	4 Display Accuracy	5 Response Time	6 Visual Quality	7 Functionality	8 Projection Range	9 Calibration Autonomy	10 Recovery Time
	Direction of Improvement		▼	▼	▲	▲	▲	▲	▲	▲	▲	▲
1	Ease of Use	▲	•	•	○	○	○	△	•	○	△	○
2	Cost	▼	•	•	•	•	•	•	•	•	•	○
3	Latency	▼	△	△	•	○	•	△	•	•	△	•
4	Connectivity	▲	△	△	△	•	•	•	○	•	△	○
5	User Safety	▲	○	○	△	△	△	△	△	△	△	△
6	System Display	▲	△	○	○	•	△	•	•	△	△	△
7	Portability	▲	•	•	△	△	△	△	△	△	○	△
8	Communication Range	▲	△	△	△	•	•	△	•	•	△	△
9	Ease of Maintainence	▲	•	•	○	△	△	△	•	△	•	△
10	Data Accuracy	▲	△	△	•	•	○	•	•	•	△	○
	Target		12" x 12" x 5"	8 lbs	~ 1 hr	100%	< 45 seconds	1920 x 1080 pixels	< 10% failure	> 6 ft	Fully Autonomous	< 10 seconds

Correlations:
Positive: +
Negative: -
No Correlation: ~

Relationships:
Strong = •
Medium = ○
Weak = △

Direction of Improvement:
Maximize: ▲
Target: ☐
Minimize: ▼

Figure 2.4.6 *House of Quality*

2.5 Diagrams

2.5.1 Diagram Overview

The system is composed of a projector with a camera and touch screen monitor attached, as well as hardware to allow onboard calibration of the image to the target surface. The overall system can be displayed via the Hardware and Software Diagrams listed in Figure 2.5.2 and Figure 2.5.3, respectively.

2.5.2 Hardware Diagram

The Hardware Diagram is listed below with a Key that indicates which team member is in charge of what aspect of the Hardware Integration. The *Key* also depicts the connection type.

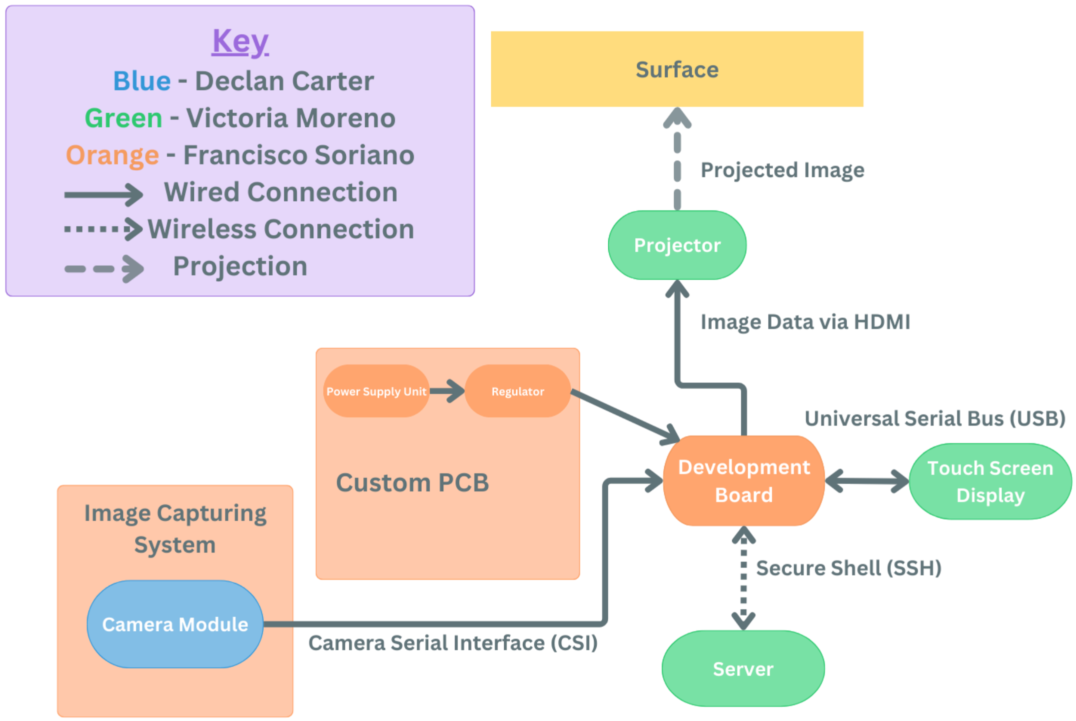


Figure 2.5.2 *Hardware Block Diagram*

A custom PCB will house a power supply unit that will supply current into a regulator in order to properly match the MCU's input voltage requirements. This custom PCB will have a wired connection with the MCU. Additionally, an image capturing system, consisting of IR Sensors and a camera module, will also have a wired connection with the MCU. A touch screen display will also have a wired connection with the MCU. Furthermore, a server will have a wireless connection with the MCU. Ultimately, the MCU will be connected to the projector which will project images onto a surface.

2.5.3 Software Diagram

The Software Diagram is listed below with a key that indicates which team member is in charge of the various aspects of the Software Integration. The Key also depicts the different shapes of blocks in the diagram that indicate what it requires.

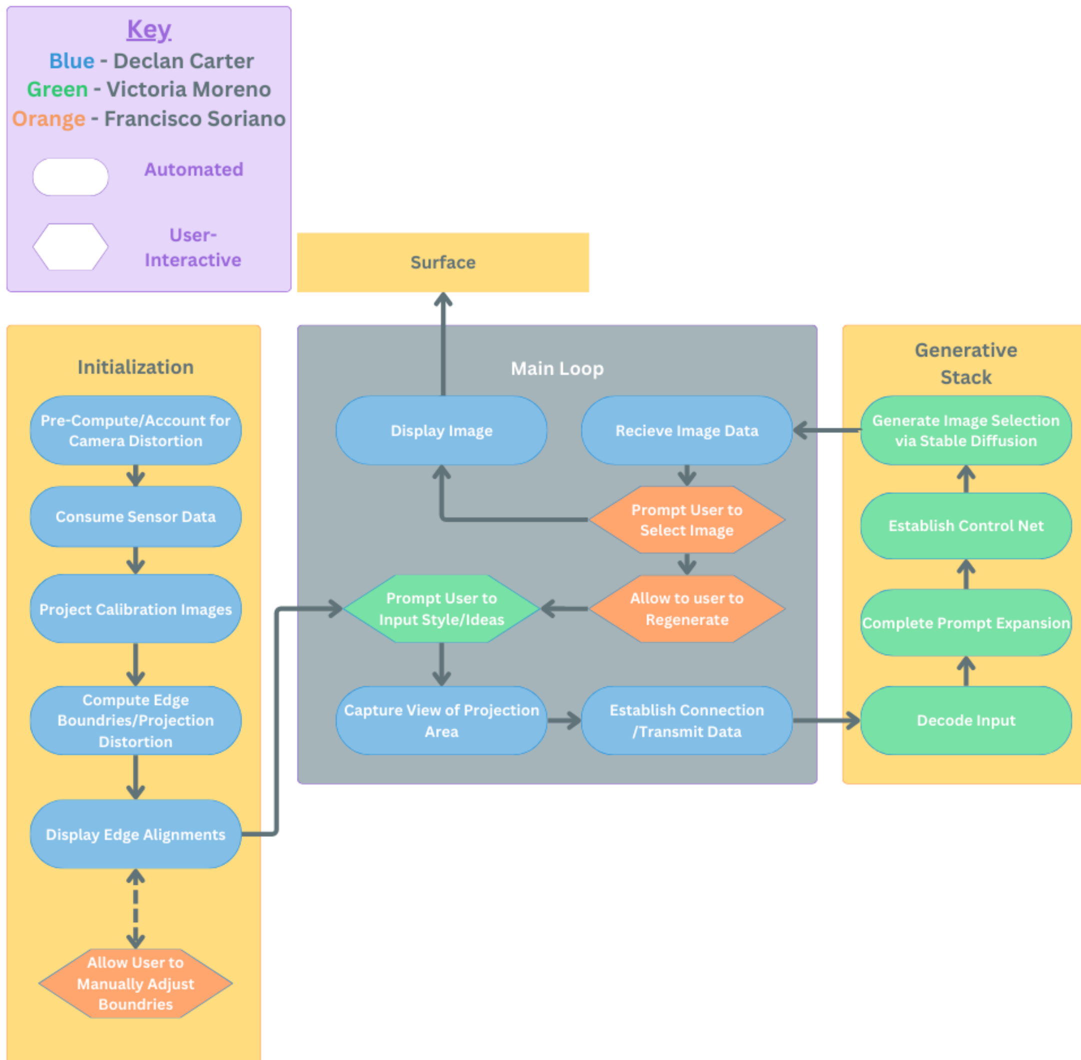


Figure 2.5.3 *Software Block Diagram*

When the device is turned on, initial calibration is needed, requiring varying levels of human interaction. To begin, lens distortion is automatically accounted for to smooth/transform the edges of the otherwise rounded image to be normal with the projected plane. Next, the device will project the Canny Edges it detects onto the surface, allowing the user to manually tune the corners of the image, aligning the projected vertices with the vertices on the wall.

Moving on to the main loop, there will exist a UI that allows the users to prompt the image generation model, which, in turn, being connected to the Generative AI stack, packages all relevant sensor data, as well as the included prompt and ships it to the off-network device, as discussed further in Section 7.2.4.

On this server, the packet is then received and decoded. However, due to the specific nature of the GPMS, further prompt expansion will be required to create an image suitable for projection. Additionally, within the image generation pipeline, ControlNet will be utilized as the primary means of ensuring the same edges detected in the image will go on to align with the vertices on the given surface. Finally, stable diffusion will be used to create a number of suitable images to be sent back to the device, allowing the user to select the image most aligned with their vision before finally being displayed on the given structure.

2.6 GPMS - An Illustration

The image below was generated using ChatGPT 4 to fuse the elements of modern technology into a single device. [1] It features a projector with two cameras, capturing multiple views. Our image is sleek and futuristic, with a silver and black color scheme common in contemporary technology devices. The ports and ventilation grills lend to high functionality and processing power, which would be necessary for high-definition projection onto a 3D surface. This image is a visual representation and is currently not available in the consumer market. It is intended to be used solely for brief illustrative purposes during technical presentations of all disciplines and to further envision future technology.

Figure 2.6 *GPMS - An Illustration*

Chapter 3 Research

3.1 Hardware Integration Strategies

In the field of projection technology, there are two primary approaches to hardware integration: a self-contained solution that provides all necessary components/hardware and a user-dependent model that relies on user-provided hardware. The user-dependent model, exemplified by products like LumaBox, requires users to download the software onto their devices, capture images of structures using their own equipment, and then transfer these images to a computer for processing. The computer, in turn, is connected to a projector to display the final output.

Alternatively, a self-contained solution bundles all necessary components - the projector, webcam, computer, and monitor - into a single, integrated product. This approach allows for more intricate low-level work, ensuring seamless communication between the components and simplifying the user experience, particularly for those who may not be technologically adept.

As a team from the College of Electrical and Computer Engineering at the University of Central Florida, the authors' focus is on creating a comprehensive hardware solution. They firmly believe that including all necessary components within the product will not only align with their academic focus but also enhance the user experience by reducing the need for technical expertise and ensuring compatibility between components.

Central to the authors' self-contained approach is the integration of a webcam mounted on the projector. This design choice means that users only need to position the projector correctly and specify when the image should be captured, simplifying the setup process and minimizing the need for additional user-provided hardware.

However, the authors also considered the potential benefits and drawbacks of user-dependent solutions, as outlined in the following table:

Aspect	User-Contained Solutions	User-Dependent Solutions
Compatibility	Optimized for seamless compatibility between components.	Requires compatibility with a wide range of user devices, potentially leading to unforeseen issues.
User Complexity	Simplifies the user experience by reducing the	Adds complexity at the software and hardware levels, requiring users to

	need for additional device configuration.	manage multiple devices and configurations.
Software Requirements	Tailored to the integrated hardware, minimizing software complexity.	Could necessitate the development of platform-agnostic applications, increasing the workload and complexity.
Implementation	Straightforward, focusing on the integrated components without adding layers of complexity in software and hardware management.	Introduces significant complexity, requiring compatibility management for various user-provided components.
Scalability	Limited scalability due to the fixed configuration of integrated components.	Highly scalable, as users can upgrade their own hardware components as needed.

Table 3.1 *User-Contained vs User-Dependent Solutions*

After careful consideration, the authors determined that the benefits of a self-contained solution outweigh the potential advantages of relying on user-dependent solutions. By providing a complete, fully integrated product, the authors can ensure a streamlined, user-friendly experience while minimizing compatibility issues and software complexity. This approach aligns with their academic focus, simplifies the user experience, and sets their project apart by offering a superior, turnkey solution that optimizes performance through carefully selected and integrated components.

3.2 Hardware Comparison and Selection

3.2.1 MCU vs FPGA

The authors were torn at first when deciding whether they should proceed with an MCU or an FPGA. After talking with various faculty mentors and discussing their experiences with the two options, the authors came to the following conclusions.

In deciding between an MCU and a field-programmable gate array (FPGA) for GPMS, there were distinct trade-offs that were heavily considered. While FPGAs offer superior processing power and precision tailored for specialized tasks, they come with higher costs, increased power consumption, and complexity in programming. Conversely, to the benefit of GPMS, MCUs prioritize low power consumption and user-friendliness, featuring readily available peripherals, extensive documentation, and support for common programming languages, simplifying development and debugging processes. [9]

The choice between an FPGA and MCU hinges on factors such as power efficiency, ease of use, and scalability which will be further explored throughout the research stage of the project. MCUs offer a straightforward development process with built-in features for stability and standardized architecture, making them ideal for rapid prototyping and easy integration. In contrast, FPGAs actually demand specialized knowledge and customized solutions for error handling, yet they provide unparalleled scalability for future system upgrades. Additionally, while MCUs boast environmental robustness with built-in temperature sensors and watchdog timers, FPGAs may require additional measures to mitigate sensitivity to environmental factors like temperature and electromagnetic interference. [9]

Opting for an MCU over an FPGA can offer distinct advantages in specific project scenarios, particularly those requiring a balance of computational power, versatility, and ease of use. The table below encompasses a direct comparison between an MCU and an FPGA. [9]

Feature	MCU	FPGA
Computational Power	Sufficient for a wide range of tasks, including multimedia processing, networking, and computing	Superior, tailored for specialized tasks
Power Consumption	Low, prioritizing energy efficiency	Higher, due to enhanced processing capabilities
Programming Complexity	User-friendly; supports common programming languages, simplifying development and debugging	Complex; requires specialized knowledge for programming and debugging
Development Process	Straightforward with built-in features for stability, extensive documentation, and support	Demands customized solutions for error handling, scalability
Scalability	Standardized architecture, ideal for rapid prototyping and integration	Unparalleled, offers flexibility for future system upgrades

Environmental Robustness	Built-in sensors and timers for temperature and stability	May require additional measures for environmental factors
Operating System	Linux-based, providing a familiar environment with extensive software support	Not specified; often requires custom software solutions
Connectivity and Peripherals	Extensive, including USB, HDMI, Ethernet, Wi-Fi, Bluetooth	Not specified; customization might be necessary for specific needs
Cost	Generally lower, making it affordable and accessible	Higher, reflecting the advanced capabilities and flexibility

Table 3.2.1 *MCU vs FPGA*

After careful consideration, the authors have opted to use an MCU. Overall, the combination of computational power, software compatibility, connectivity options, and affordability make it a compelling choice for GPMS requiring a flexible and user-friendly MCU solution, particularly when compared to the specialized and potentially more complex nature of alternatives such as FPGAs. [9]

The MCU presents several compelling reasons for its suitability in the context of GPMS such as providing sufficient computational capabilities for various tasks, including multimedia processing, networking, and general-purpose computing. [9]

Additionally, the MCU can operate on different operating systems (OS), which provides a familiar computing environment and extensive software support. This compatibility with different operating systems enables developers to leverage a vast ecosystem of software libraries, tools, and applications, facilitating rapid prototyping and development. [9]

Moreover, the MCU offers an opportunity to include a rich set of built-in peripherals and connectivity options, including USB ports, HDMI output, Ethernet, Wi-Fi, and Bluetooth. These features make it highly versatile, allowing for seamless integration with various devices and systems, such as sensors, displays, and network interfaces. [9]

Overall, the combination of computational power, software compatibility, connectivity options, and affordability make the MCU a compelling choice for GPMS requiring a flexible and user-friendly solution, particularly when compared to the specialized and potentially the more complex nature of alternatives such as FPGAs. [9]

3.2.2 MCU vs Development Board

When moving forward with the development of the project, the authors explored the possibilities of using either a development board or an MCU. Therefore, the authors explored a few aspects surrounding the two including Hardware Integration, Processing Power, Rich Peripheral Support, Software Development Environment, and Community Support and Resources.

Development boards often come with integrated components such as high-quality cameras. This integration reduces the complexity of sourcing and interfacing separate hardware components, ensuring compatibility and streamlined development.

Creating an immersive storytelling environment with dynamic image generation and sophisticated computer vision techniques requires significant processing power. Development boards typically feature more powerful processors than simple MCUs, allowing for faster image processing, AI inference, and real-time interactions with all users.

The project requires interaction with various peripherals such as cameras, touchscreens, and projectors. Development boards offer a wide range of peripheral interfaces and support for protocols like SPI, I2C, and HDMI, making it easier to connect and communicate with external devices.

Additionally, development boards often come with comprehensive software development kits (SDKs) and libraries tailored to the specific hardware features. These SDKs provide APIs and tools for rapid prototyping, debugging, and optimizing the software, accelerating the development process and reducing time-to-market.

Notably, development boards, especially popular ones, benefit from a large community of developers and enthusiasts. This community provides valuable resources such as documentation, tutorials, sample code, and forums for troubleshooting and sharing knowledge, easing the learning curve and fostering innovation.

A development board offers several advantages over a simple MCU for the goals outlined in the project which can be highlighted in the table below.

Aspect	**Development Board**	**MCU**
Processing Power	Typically higher, often featuring multicore	Limited processing power, single-core

	processors	
Integrated Peripherals	Comprehensive, including sensors, displays, etc.	Basic, may require additional components
Connectivity Options	Extensive, including Wi-Fi, Bluetooth, Ethernet, etc.	Limited, may require external modules
Development Environment	Rich software development kits (SDKs) and libraries	Basic toolchains, limited development support
Community Support	Large community with extensive resources	Limited community, fewer resources
Cost	Higher due to integrated features and performance	Lower, especially for simple MCU setups

Table 3.2.2 *Development Board vs MCU Comparison*

Development boards offer a comprehensive solution for the GPMS project, integrating essential hardware components such as high-quality cameras, touchscreens, and powerful projectors. This integration effectively streamlines hardware setup and also ensures compatibility, simplifying the development process.

In addition to integrated hardware, development boards provide superior processing power compared to simple MCUs. This enhanced capability enables faster image processing and AI inference, crucial for creating an immersive storytelling environment with dynamic image generation.

Furthermore, development boards offer comprehensive peripheral support, reducing the complexity of interfacing external devices like cameras and touchscreens. Coupled with robust software development kits (SDKs) and libraries tailored to their hardware features, development boards facilitate efficient software development, debugging, and optimization. Supported by a large community of developers and enthusiasts, development boards provide valuable resources and support, making them the preferred choice for efficiently realizing the ambitious goals of the GPMS project. [10] [11]

3.2.3 Development Board Comparison

After narrowing down the options, the authors compared a Raspberry Pi 4B, ESP32-WROOM-32, and MSP-EXP430fr6989. Each of the development boards are briefly, yet comprehensively, listed below in the respective tables.

The **Raspberry Pi 4B** has the following stand out features to integrate with the custom PCB: [12] [13]

- Broadcom BCM2711, quad-core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz
- 1GB, 2GB, 4GB, or 8GB LPDDR4 (varies by model) with on-die ECC
- 2.4 GHz and 5.0 GHz IEEE 802.11b/g/n/ac wireless LAN, Bluetooth 5.0, BLE, Gigabit Ethernet, 2 × USB 3.0 ports, 2 × USB 2.0 ports
- Standard 40-pin GPIO header (fully backwards-compatible with previous boards)
- 2 × micro HDMI ports (supports up to 4Kp60), 2-lane MIPI DSI display port, 2-lane MIPI CSI camera port
- 4-pole stereo audio and composite video port
- H.265 (4Kp60 decode), H.264 (1080p60 decode, 1080p30 encode), OpenGL ES, 3.0 graphics
- Micro SD card slot for loading operating system and data storage
- 5V DC via USB-C connector (minimum 3A), **5V DC via GPIO header (minimum 3A)**, Power over Ethernet (PoE)–enabled (requires separate PoE HAT)
- Operating temperature 0–50ºC

The **ESP32-WROOM-32** has the following stand out features to integrate with the custom PCB: [14]

- Wi-Fi Protocols
  - 802.11 b/g/n (802.11n up to 150 Mbps)
  - A-MPDU and A-MSDU aggregation with 0.4 µs guard interval support
  - Center frequency range of operating channel: 2412 ~ 2484 MHz
- Bluetooth
  - Protocols: Bluetooth v4.2 BR/EDR and Bluetooth LE specification
- Radio
  - NZIF receiver with sensitivity of -97 dBm
  - Class-1, class-2, and class-3 transmitter
  - Adaptive Frequency Hopping (AFH)
- Hardware:
  - Module interfaces: SD card, UART, SPI, SDIO, I2C, LED PWM, Motor PWM, I2S, IR, pulse counter, 39 GPIO, capacitive touch sensor, ADC, DAC, Two-Wire Automotive Interface (TWAI®)
  - Integrated crystal: 40 MHz crystal

    - Integrated SPI flash: 4 MB
    - Operating voltage/Power supply: 3.0 V ~ 3.6 V
    - Operating current (Average): 80 mA
    - Minimum current delivered by power supply: 500 mA
    - Recommended operating ambient temperature range: -40 °C ~ +85 °C
    - Package size: 18 mm × 25.5 mm × 3.10 mm
    - Moisture sensitivity level (MSL): Level 3

The **MSP-EXP430fr6989** has the following stand out features to integrate with the custom PCB: [15]

- MSP ULP FRAM-based MSP430FR6989 16-bit MCU
    - 100 uA/MHz active mode and 350 nA standby with RTC and 3.7 pF crystal
    - Certified ULPBench score of 109
    - 128 KB FRAM
    - 16-Bit RISC architecture up to 8-MHz FRAM access/ 16MHz system clock speed
    - 320-segment LCD controller
    - Extended Scan Interface
    - 16 channel 12-bit ADC
    - Comparator
    - 5 Timers
    - Direct memory access
    - 256-bit AES
    - 83 GPIO
- EnergyTrace++™ Technology available for ultra-low-power debugging
- 40 pin LaunchPad standard leveraging the BoosterPack ecosystem
- Onboard eZ-FET emulation
- 2 buttons and 2 LEDs for User Interaction
- Segmented LCD
- Pins for direct access to the Extended Scan Interface

It is important to note that the overall project deliverable will benefit most from the development board, that produces higher processing power, contains greater memory capacity, allows for direct video output, provides an operating system and multimedia support, and hosts various peripheral options. The comparisons of these factors are listed in the table below.

Main Factors	**Raspberry Pi 4B**	**ESP32-WROOM-32**	**MSP-EXP430fr6989**
Processor	Broadcom BCM2711 quad-core	Single/Dual-Core 32-bit LX6 Microprocessor	MSP ULP FRAM-based MSP430FR6989

	Cortex-A72 64-bit SoC @ 1.5GHz	up to 240 MHz	16-bit MCU
Memory	1GB, 2GB, 4GB, or 8GB LPDDR4	4 MB SPI flash	128 KB FRAM
Wireless Connectivity	2.4 GHz and 5.0 GHz IEEE 802.11b/g/n/ac wireless LAN, Bluetooth 5.0, BLE	802.11 b/g/n (802.11n up to 150 Mbps), Bluetooth v4.2 BR/EDR and Bluetooth LE specification	N/A
Ports	2 × USB 3.0 ports, 2 × USB 2.0 ports	N/A	N/A
GPIO	Standard 40-pin GPIO header (fully backwards-compatible with previous boards)	39 GPIO, capacitive touch sensor	83 GPIO
Video Output	2 × micro HDMI ports (supports up to 4Kp60), 2-lane MIPI DSI display port, 2-lane MIPI CSI camera port, 4-pole stereo audio and composite video port	None	N/A
Operating System and Multimedia Support	H.265 (4Kp60 decode), H.264 (1080p60 decode, 1080p30 encode), OpenGL ES, 3.0 graphics	N/A	N/A
Power Supply	5V DC via USB-C connector (minimum 3A), 5V DC via GPIO	Operating voltage/Power supply: 3.0 V ~ 3.6 V	N/A

	header (minimum 3A), Power over Ethernet (PoE)–enabled (requires separate PoE HAT)		
Operating Temperature	0–50ºC	-40 °C ~ +85 °C	N/A

Table 3.2.3 *Development Board Comparison*

In essence, the Raspberry Pi 4B is most powerful but consumes the most power. The Raspberry Pi 4B is the optimal solution for the project because it has higher processing power, greater memory capacity, direct video output, operating system and multimedia support, and various peripheral options. These features make it more suitable for complex projects that require high computational power and versatility. It also has a large community and extensive documentation, which can be very helpful during the development process.

In summary, despite the ESP32-32-WROOM-32 and MSP-EXP430fr6989 being two versatile and powerful microcontrollers for a select range of IoT applications, it does not suit the GPMS sufficiently. On the other hand, the Raspberry Pi 4B lends itself to me as the more suitable MCU for GPMS.

3.2.4 Projector Comparison

The tables below encapsulates the technical comparison of technology, color production, brightness, throw ratio, suitability for well-lit rooms, features for educational and business environments, and an overall technical and holistic recommendation for GPMS.

The projector that will be used for GPMS serves as an essential part of the entire system. Francisco initially owned an Epson PowerLite 585W ultra short throw projector. Therefore, before the comparison the authors decided to test the functionality of the projector already owned and available. After testing its functionality, the authors decided collectively not to use it.

The decision not to use the ultra throw projector, specifically the Epson PowerLite 585W, stems from considerations regarding image distortion and the practicality of correcting it. When projecting onto a surface, especially in unconventional settings like corners or non-flat surfaces, there can be significant image warping. The distance and angle at which the projector emits light relative to the surface can cause distortion. In this case, the positioning of the Epson PowerLite 585W might result in warping that could distort the image, particularly in the corners.

The authors have determined that the camera's angle relative to the surface is crucial for capturing images with minimal distortion. By ensuring that the camera angle is perpendicular to the surface, they aim to capture an even image with as little distortion as possible. This suggests a deliberate effort to optimize the setup for image clarity and accuracy.

To counteract the distortion caused by the projector's angle and the surface's geometry, a corner transformation would be necessary. This process involves digitally altering the image to correct for the distortion caused by the unique projection setup. However, implementing corner transformation can be complex and may not always yield satisfactory results, especially if the distortion is severe. Furthermore, adjusting the image through warping and transforms can be challenging, requiring precise calibration and potentially leading to image degradation or artifacts.

Considering these factors, the authors concluded that using a different projector better suited to the setup would be more practical and efficient. This alternative projector may offer features or specifications that align more closely with the desired imaging requirements, minimizing the need for extensive digital correction or compromising on image quality. Ultimately, the goal of GPMS is to achieve the best possible immersive image projection with minimal distortion and technical complications.

Technical Comparison of Projectors is as follows. [16] [17] [18] [19]

Feature	**M8-F Auking**	**HAPPRUN H1**	**Panseba**
Price (USD)	57.98	79.99	99.99
Brightness (ANSI)	136	249	350
Native Resolution	1920 x 1080	1920 x 1080	1920 x 1080
Display Resolution	1080p	1080p	1080p
Connectivity	HDMI/USB/VGA	HDMI/USB	HDMI/USB
Maximum Image Size (Inches)	200	200	300
LED Life (Hours)	55,000	100,000	100,000
WiFi Connection	No	No	Yes

Bluetooth Connection	No	Yes	Yes
Weight (Pounds)	2.87	3.4	3.3
Dimensions (Inches)	7.95 x 6.3 x 2.91	9.3 x 8.1 x 3.5	9.0 x 7.0 x 3.0
Contrast Ratio	2000:1	10,000:1	10,000:1
Max Throw Distance (Feet)	18.5	22.97	16.4

Table 3.2.4 *Technical Projector Comparison*

The authors collectively decided to select the Panseba projector. Given the comparison in Table 3.2.4, the Panseba projector stands out for several reasons. The Panseba projector offers several notable features setting it apart from competitors such as the M8-F Auking and HAPPRUN H1.

With a brightness of 350 ANSI, the Panseba projector stands out as significantly brighter than both the M8-F Auking and HAPPRUN H1. This brightness ensures a clearer and more vibrant image, particularly in well-lit environments.

In terms of maximum image size, the Panseba projector can project up to 300 inches, surpassing both the M8-F Auking and HAPPRUN H1. This larger projection size enhances the immersive viewing experience.

One distinct advantage of the Panseba projector is its WiFi connection, a feature lacking in the other two projectors. This WiFi connectivity provides greater flexibility for content streaming and device connectivity.

Additionally, the Panseba projector boasts an impressive LED life of 100,000 hours, equaling that of the HAPPRUN H1 but surpassing the M8-F Auking by a considerable margin. This extended LED life means the projector can operate for longer periods before requiring a bulb replacement. The extent of the LED life that the Panseba projector offers is imperative for GPMS to operate for a long period of time.

For seamless connectivity with Bluetooth devices such as speakers or headphones, the Panseba projector offers a Bluetooth connection, akin to the HAPPRUN H1.

In addition to its impressive features, the Panseba projector's compact size further enhances its appeal, making it an excellent choice for portability.

Measuring just 9 by 7 by 3 inches, this projector is compact and lightweight, making it easy to transport from one location to another.

The compact dimensions of the Panseba projector do not compromise its performance or features. Despite its small size, it still offers impressive brightness, maximum image size, connectivity options like WiFi and Bluetooth, and a long LED life. This combination of portability and functionality makes the Panseba projector a versatile solution for various applications specific to those of GPMS.

Finally, the Panseba projector's contrast ratio of 10,000:1 matches that of the HAPPRUN H1 but exceeds the M8-F Auking's. A higher contrast ratio translates to superior color depth, resulting in a more dynamic and vibrant image quality.

While the Panseba projector is more expensive, these features justify the higher price point and make it a superior choice based on the table provided. Given that portability is a priority, the lighter weight of the M8-F Auking might be an advantage.

3.2.5 Camera Comparison

Selecting the appropriate camera type is crucial for the optimal computer vision integration and quality of the GPMS device. The camera is integral to the functionality of the system, capturing high-quality images necessary for the software to perform accurate analyses. Therefore, the camera must not only be compatible with the Raspberry Pi 4B but also meet specific performance criteria, including resolution, frame rate, and low-light capabilities. This section will compare various camera types, evaluating their suitability based on these essential parameters to ensure the chosen camera enhances the effectiveness and efficiency of the GPMS device.

For the project, 2 distinct camera data interfaces are to be compared: Universal Serial Bus (USB) and Camera Serial Interface (CSI). The following table compares these interfaces with the features they provide [20]:

Feature	USB Camera	CSI Camera
Compatibility	High compatibility with many devices	Primarily compatible with Raspberry Pi models
Connection	Connects via USB port	Connects through a dedicated CSI port

Ease of Setup	Generally plug-and-play	Requires specific setup on Raspberry Pi
Data Transfer Speed	Slower compared to CSI; depends on USB type	Faster, designed for high data rates
Power Consumption	Typically higher	Generally lower
Flexibility	More flexibility in positioning	Fixed positioning close to the Raspberry Pi
Image Quality	Varies widely; generally lower than CSI	Typically higher, optimized for image capture
Cost	Can be less expensive	Often more expensive but offers better quality

Table 3.2.5.1 *Camera Interface Comparison*

The comparison between USB cameras and CSI cameras reveals distinct advantages and trade-offs for each type. USB cameras offer high compatibility with various devices and are generally easier to set up due to their plug-and-play nature. However, they may suffer from slower data transfer speeds and higher power consumption compared to CSI cameras. On the other hand, CSI cameras, while primarily compatible with Raspberry Pi models and requiring a more specific setup, provide faster data transfer rates and superior image quality. These cameras are also more energy-efficient but tend to be more expensive.

However, taking into account the above features, the authors have decided that a higher quality image and smaller power footprint than a standard USB camera makes a camera featuring CSI a straightforward decision for the project.

To enhance the capabilities of the GPMS device, incorporating a high-quality camera using the Camera Serial Interface (CSI) is imperative. This section focuses on two prominent CSI cameras: the OV5640 and the IMX477. Each camera brings unique advantages to image capturing tasks, tailored to different performance needs and technical requirements. Below is a comparative overview of their key features.

Feature	OV5640	IMX477
Resolution	5 Megapixels (2592x1944)	12.3 Megapixels (4056x3040)
Sensor Type	CMOS	CMOS
Frame Rate	15 fps at full resolution	60 fps at full resolution
Lens Compatibility	Fixed lens	Interchangeable lenses supported
Video Output	Supports 1080p at 30 fps	Supports 4K at 30 fps
Price Range	Budget-friendly	Premium pricing

Table 3.2.5.2 *Overall Camera Comparison*

The OV5640 and IMX477 cameras offer robust solutions for projects requiring high-resolution imaging through the CSI interface. The OV5640 is a budget-friendly option suitable for applications requiring standard resolution and frame rates, while the IMX477 caters to more advanced needs with its higher resolution, faster frame rates, and support for interchangeable lenses.

However, it’s important to note the application at hand and the limited computing power of the Raspberry Pi, specifically when looking to process a video feed at 4k 30 fps. While the resolution could easily be downscaled, the quality offered by the IMX477 seems to be overboard, especially on a project whose cost is an important factor. These reasons have led the authors to ultimately choose the OV5640.

3.2.6 Monitor Comparison

To ensure a seamless user experience and maintain the compact nature of the GPMS device, it is essential to integrate a monitor that connects directly to the Raspberry Pi and serves as the frontend interface. The primary requirement for this monitor is that it must be a touch screen, eliminating the need for additional peripherals such as a mouse and keyboard. A touchscreen monitor allows for intuitive navigation and interaction with the GPMS software, streamlining the user experience and reducing the overall footprint of the device.

When selecting a touchscreen monitor for the GPMS device, several factors should be considered, including screen size, resolution, touch technology, compatibility, and cost. The following chart compares the three main types of touchscreen monitors: [21]

Feature	Resistive Touch Screen	Capacitive Touch Screen	Infrared Touch Screen
Touch Technology	Pressure-sensitive; works with any object	Responds to electrical conductivity; works with fingers or capacitive stylus	Uses infrared light to detect touch; works with any object
Durability	Can withstand scratches and impacts; suitable for harsh environments	More prone to scratches and damage; not suitable for harsh environments	High durability; can withstand scratches and impacts
Accuracy	Less precise; may require calibration	High precision and responsiveness	High precision and responsiveness
Multi-touch Support	Limited or no multi-touch support	Excellent multi-touch support	Good multi-touch support
Visibility	May have reduced clarity due to additional layers	Clear display; no additional layers affecting visibility	Clear display; no additional layers affecting visibility
Cost	Generally less expensive	More expensive than resistive touch screens	Often the most expensive option

Table 3.2.6.1 *Monitor Overall Comparison*

Based on the requirements of the GPMS device, a capacitive touch screen monitor would be the most suitable choice. Capacitive touch screens offer high precision, excellent multi-touch support, and a clear display, which are all essential for a user-friendly interface. Although they may be more expensive than resistive touch screens, the benefits they provide in terms of user experience and functionality justify the investment. [21]

When making the final decision, it is important to consider the specific Raspberry Pi model being used and ensure that the chosen touch screen monitor is compatible with both the Raspberry Pi and the GPMS software. Additionally, the screen size should be appropriate for the intended use case, balancing readability and portability.

There are many different brands and types of capacitive touch screens. The authors decided that the best size would be 7 to 10 inches measured by its diagonal. This is because it would not be too big, increasing the cost and making the device bulkier, but not too small where the interface is hard to see. Now the authors need to decide which size to pick. The chart below compares the two.

Factor	**7-inch Screen**	**10-inch Screen**
Screen Size	Smaller screen size may make it more challenging to design a user-friendly interface that displays all necessary information.	Larger screen size provides more space to create a user-friendly interface that can display all required information without excessive scrolling or cluttering.
Typing Experience	Smaller touch targets for on-screen typing may lead to more user errors and a less comfortable typing experience.	Larger touch targets for on-screen typing can improve user accuracy and provide a more comfortable typing experience, reducing the likelihood of errors.
Calibration Precision	Users may encounter difficulties accurately dragging the corners of an image to the desired locations due to the smaller screen size.	The larger screen size allows for more precise dragging of image corners during calibration, potentially resulting in better calibration accuracy and a smoother user experience.
Information Display	Limited screen real estate may necessitate the use of scrolling, multiple pages, or smaller font sizes to display all information.	More screen space enables the display of more information at once, reducing the need for scrolling or multiple pages and allowing for larger, more readable font sizes.
UI Design Complexity	Designing a user interface that effectively accommodates all required elements and information may be more challenging.	The larger screen size provides more flexibility in UI design, making it easier to create an intuitive and visually appealing interface that incorporates all necessary components without appearing cluttered.

Portability	7-inch screens are more portable and easier to integrate into compact devices, which may be advantageous for certain use cases.	10-inch screens are less portable and may result in a larger overall device size, which could be a drawback if portability is a key concern.
Cost	7-inch screens are generally more affordable compared to larger screens.	10-inch screens are typically more expensive than 7-inch screens due to the larger display size and potentially higher resolution.

Table 3.2.6.2 *Monitor Size Comparison*

After carefully considering the advantages and disadvantages of both 7-inch and 10-inch capacitive touch screens, the authors have concluded that a 10-inch monitor would be the most suitable choice for the GPMS device. While a 7-inch screen might offer better portability and lower costs, the benefits of a larger 10-inch display outweigh these factors in terms of user experience and functionality.

A 10-inch screen provides ample space to create a user-friendly interface that can display all necessary information without excessive scrolling or cluttering. This larger screen size allows for more precise dragging of image corners during calibration, potentially resulting in better calibration accuracy and a smoother user experience. Additionally, the increased screen real estate enables the display of more information at once, reducing the need for scrolling or multiple pages and allowing for larger, more readable font sizes.

Furthermore, the 10-inch screen offers more flexibility in UI design, making it easier to create an intuitive and visually appealing interface that incorporates all necessary components without appearing cluttered. The larger touch targets for on-screen typing can also improve user accuracy and provide a more comfortable typing experience, reducing the likelihood of errors.

While a 10-inch screen may result in a slightly larger overall device size and higher costs compared to a 7-inch screen, the authors believe that the enhanced user experience and functionality justify these trade-offs. The improved precision, readability, and design flexibility offered by a 10-inch display will ultimately lead to a more effective and user-friendly GPMS device.

In conclusion, the authors have determined that a 10-inch capacitive touch screen monitor is the optimal choice for the GPMS device, as it strikes the perfect balance between screen size, user experience, and functionality. This decision will ensure that the device meets the requirements of the intended use case while providing a high-quality, user-friendly interface that enhances the overall effectiveness of the GPMS system. Now here are 3 options for 10-inch capacitive touch screens.

Feature	GeeekPi 10.1” Capacitive Touchscreen	Waveshare 10.1inch HDMI LCD (H) with Capacitive Touch	Elecrow 10.1 Inch 1920x1080 HDMI TFT LCD Display with Touch Screen
Screen Size (diagonal)	10.1 inches	10.1 inches	10.1 inches
Resolution	1024 x 600 pixels	1024 x 600 pixels	1920 x 1080 pixels
Touch Technology	Capacitive multi-touch	Capacitive multi-touch	Capacitive multi-touch
Connection Interface	HDMI for display, USB for touch	HDMI for display, USB for touch	HDMI for display, USB for touch
Compatibility	Compatible with Raspberry Pi 4B	Compatible with Raspberry Pi 4B	Compatible with Raspberry Pi 4B
Dimensions	235mm x 142mm x 8mm	239mm x 169mm x 30mm	229mm x 149mm x 23mm
Power Supply	Powered via USB connection	Requires external power supply	Requires external power supply
Viewing Angle	178°(H) / 178°(V)	170°(H) / 170°(V)	85°(H) / 85°(V)
Price Range	\$80 - \$90	\$100 - \$120	\$120 - \$140

Table 3.2.6.3 *Selected Monitor Specifications Comparison*

After careful consideration of the three 10-inch capacitive touch screen options, the authors have decided to proceed with the GeeekPi 10.1” Capacitive Touchscreen for the GPMS device. This decision was based on several key factors that align with the project's requirements and priorities. Firstly, the GeekPi display offers a high resolution of 1024 x 600 pixels, which ensures a clear and visually appealing user interface. Secondly, its wide viewing angles of 178° both horizontally and vertically allow for excellent visibility from various positions, making it suitable for use in different environments.

Additionally, the GeekPi display is compatible with multiple Raspberry Pi models, providing flexibility and future-proofing for the project. Most importantly, its power comes via USB, meaning it can be powered using only the existing Power Supply Unit (PSU), which effectively powers the entire system by connecting to the Raspberry Pi. Lastly, its affordable price point makes it a cost-effective solution without compromising on essential features. While the other options have their strengths, such as the Elecrow display's higher resolution and the Waveshare display's wide viewing angles, the GeeekPi 10.1" Capacitive Touchscreen offers the best balance of specifications, compatibility, and affordability, making it the ideal choice for the GPMS device.

3.2.7 Generative AI Platforms

To effectively deploy AI models like Stable Diffusion, which demand high-quality GPU resources, the authors initially turned to Google Colaboratory due to its accessibility and cost-effectiveness. This cloud-based platform offers users a browser-based interface to develop and execute Python code. While Colab offers various plans, the authors' underlying aim was to utilize its free option to save money. However, the authors quickly discovered that Stable Diffusion was heavily restricted on the free tier.

Subsequently, the authors opted for Colab Pro, priced at $10/month, which grants access to faster GPUs within a time limit. Despite this added benefit, the author's efforts to run Stable Diffusion on Colab were impeded by newly identified bugs. Faced with these challenges, the authors explored alternative solutions, including procuring a more affordable GPU.

Fortunately, Advanced Micro Devices (AMD) provided the authors with access to a powerful computer directly from AMD. Recognizing the advantages of this new and improved setup, the authors decided to leverage remote access to this resource, enabling them to execute and test Stable Diffusion directly from their laptops without incurring additional expenses.

The specifications of the computer are as follows:

Component	Part Description
CPU	AMD Ryzen 9 7950X - 16-Core 4.5 GHz - Socket AM5 - 170W Processor

CPU Cooler	Noctua NH-D15 chromax.black 82.52 CFM CPU Cooler
GPU	SAPPHIRE NITRO Radeon RX 7900 XTX 24GB GDDR6 PCI Express 4.0 x16 ATX
MotherBoard	ASRock X670E Steel Legend AM5 ATX Motherboard
RAM	CORSAIR Vengeance RGB 96GB (2 x 48GB) DDR5 6000 Desktop Memory
Power Supply	Seasonic PRIME TX-1300, 1300W 80+ Titanium, Full Modular Power Supply
Case	Thermaltake Core P3 TG Pro ATX Mid Tower Case
Storage OS	Crucial T700 1 TB M.2-2280 PCIe 5.0 X4 NVME Solid State Drive
Storage APPS	SABRENT 2TB Rocket NVMe 4.0 Gen4 PCIe M.2

Table 3.2.7 *Computer Specifications*

This approach not only circumvents the limitations encountered with Google Colab but also allows the authors to work on the project from any location. By leveraging existing resources, the authors can streamline the workflow and focus on advancing their research without financial constraints. The GPU being used is the SAPPHIRE NITRO Radeon RX 7900 XTX 24GB GDDR6 PCI Express 4.0 x16 ATX.

3.2.8 Linear vs Switching Regulator

Linear and switching regulators are both commonly used circuits for regulating voltage. They differ in their efficiency, complexity, and ability to handle different input and output voltage relationships.

Linear regulators work by dissipating excess voltage as heat. This makes them simple and inexpensive, but also less efficient, especially when dealing with large differences between input and output voltage. They are best for applications with a small voltage difference and low current needs, where heat dissipation is not a major concern.

Switching regulators use transistors as switches to control the flow of current. This allows them to achieve much higher efficiency, even with large voltage differences. However, they are also more complex and generate electrical noise due to the switching action. Switching regulators are a good choice for applications that require high efficiency or precise voltage control, even with a wide range of input voltages. A switching regulator is a strong contender for a custom PCB, in the case of GPMS, thanks to its efficiency.

To begin, switching regulators are well regarded for power efficiency. Switching regulators shine in their ability to convert power with minimal loss. This translates to less heat generation, which is a major benefit for the tightly packed, yet simple, custom PCB. Linear regulators, on the other hand, dissipate excess voltage as heat, which can become a major problem in the custom PCB's space-constrained design. In the authors' custom PCB, efficiency matters. Heat generation is a major concern in their tightly packed custom PCB. Linear regulators, by their nature, lose energy as heat during voltage conversion. This wasted energy can become a problem thermally, potentially affecting other components and/or requiring additional heat mitigation strategies. Switching regulators, on the other hand, excel at converting power with minimal loss, generating less heat and simplifying thermal management in the design.

There is also the added benefit of wide range conversion. Unlike linear regulators, switching regulators can actually handle both stepping down (buck converter) and stepping up (boost converter) voltage which ultimately offers more flexibility.

Since the custom PCB is battery-powered, maximizing efficiency is crucial. Since switching regulators waste less power as heat, they can significantly extend battery life compared to linear regulators. This becomes especially important for portable devices or those needing to operate for extended periods on battery power such as GPMS.

However, switching regulators do come with some drawbacks to consider such as complexity and noise. The design and layout of a switching regulator circuit is more intricate compared to a linear regulator which certainly can add complexity to the PCB design.

Additionally, the switching mechanism in switching regulators can introduce electrical noise into your circuit. This might require additional filtering components to ensure it does not interfere with other sensitive parts of the custom PCB. So, while switching regulators offer clear advantages in efficiency and versatility, it was important for the authors to consider the trade-offs in design complexity and potential noise generation before they made their choice for the custom PCB.

However, in the authors' case, the benefits of switching regulators outweigh the costs for custom PCBs, especially when considering factors like efficiency, versatility, and overall design functionality.

Features	**Linear**	**Switching**
Efficiency	Less efficient, dissipate	More efficient, convert

	excess voltage as heat	power with minimal loss
Complexity	Simple and inexpensive	More complex, intricate design and layout
Voltage Handling	Best for small voltage difference and low current needs	Can handle both stepping down (buck converter) and stepping up (boost converter) voltage
Heat Generation	Generate more heat, can be a problem in space-constrained designs	Generate less heat, beneficial for tightly packed designs
Noise Generation	No electrical noise	Can introduce electrical noise, might require additional filtering components
Battery Life	Less efficient, shorter battery life	More efficient, can significantly extend battery life
Design Considerations	Less complex, easier to design	More complex, requires careful design and potential noise mitigation

Table 3.2.8 *Linear vs Switching Regulator*

In summary, the benefits outweigh the costs when thermal management is a concern due to space constraints and when a design requires a wider range of voltage conversion options. Extending battery life is critical for the custom PCB's role in GPMS. [22] [23] [24]

3.2.9 Specific Regulator Comparison

The design considerations surrounding three different types of regulators, DC/DC converters, that will greatly aid in the regulation of the voltage within the custom PCB, are listed in table below. [25]

Design	TPS61022	TPS61288	TPS61236P
Design Constraint	8A boost converter with	18V, 15A synchronous	8-A valley current, adjustable output

	0.5V ultra low-input voltage	boost converter	voltage synchronous boost converters
BOM Area	72mm2	515mm2	N/A
BOM Cost	$1.67 | 1ku	$3.03 | 1ku	N/A
BOM Count	7	13	13
Efficiency	91.3%	94.2%	92.7%
Frequency	1.06 MHz	500 kHz	1 MHz
Topology	Boost Passthrough	Boost	Boost
Integrated Circuit Operating Temperature	160.7 ℉	113.702 ℉	112.424 ℉
V_{out} **Peak-to-Peak**	40.53mV	49.83mV	156.23mV
V_{in} **Min**	0.5V	2V	2.3V
V_{in} **Max**	5.5V	18V	5V
V_{out} **Min**	2.2V	4.5V	2.8V
V_{out} **Max**	5.5V	18V	5.5V
I_{out} **Max**	4A	5A	7.8A

Table 3.2.9.1 *Specific Regulator Comparison*

A comprehensive comparison of the three potential battery types are as follows with an overview of the respective pros and cons. [25]

TPS61022

Pros	Cons
Lowest BOM cost: At $1.67 per 1000	Lowest output current: The TPS61022

units, the TPS61022 has the lowest cost among the three options.	can only deliver a maximum output current of 4A, which is lower than the other two options (5A for TPS61288 and 7.8A for TPS61236P).
Lowest BOM count: With only 7 components, the TPS61022 requires the least number of parts, which can simplify assembly and potentially reduce manufacturing costs.	Lower efficiency: The TPS61022 has an efficiency of 91.3%, which is lower than the TPS61288 (94.2%) and might lead to higher power consumption.
Meets the minimum input voltage requirement: The table specifies a requirement for an ultra-low input voltage of 0.5V, which only the TPS61022 can meet.	Higher peak-to-peak output voltage ripple: The TPS61022 has a higher output voltage ripple (40.53 mV) compared to the other options, which might need additional filtering depending on the application's sensitivity to noise.

Table 3.2.9.2 *Regulator 1 Pros & Cons*

TPS61288

Pros	Cons
Higher output current: Delivers up to 5A, which is more than the TPS61022's 4A.	Higher BOM cost: At $3.03 per 1000 units, it's nearly double the cost of TPS61022.
Higher efficiency: Achieves 94.2% efficiency, leading to lower power consumption compared to TPS61022.	Higher BOM count: Requires 13 components, making it more complex to assemble compared to TPS61022.
	Higher minimum input voltage: Requires a minimum input voltage of 2V, which is not suitable for the specified 0.5V requirement.
	Larger footprint: The table doesn't explicitly mention the size, but generally, higher component count translates to a larger footprint.

Table 3.2.9.3 *Regulator 2 Pros & Cons*

TPS61236P

Pros	Cons
Highest output current: Delivers up to 7.8A, significantly exceeding both other options.	BOM information not available: The table lacks information about BOM cost and count, making a direct comparison in this specific regard somewhat difficult.
	Higher minimum input voltage: Requires a minimum input voltage of 2.3V, exceeding the specified 0.5V requirement.
	Larger footprint (likely): Similar to TPS61288, the higher component count suggests a potentially larger footprint.

Table 3.2.9.4 *Regulator 3 Pros & Cons*

The TPS61288 and TPS61236P, while powerful and efficient, are designed for applications requiring higher currents and voltages than what the Raspberry Pi 4B typically needs. Their higher cost, larger size, and the complexity of the design are not necessary for powering the Raspberry Pi 4B and therefore not considered further.

Therefore, since GPMS prioritizes low cost and low component count and can operate within the 4A current limit, the TPS61022 makes itself to be the best and simplest choice. Choosing the correct regulator with minimal connections was vital to a functioning custom PCB to ensure proper voltage regulation. Regulators ensure that the voltage supplied to various components on the PCB remains constant. This stability is especially crucial for the proper operation of sensitive electronic components housed in the custom PCB, preventing damage from overvoltage conditions and ensuring consistent performance. Striking the balance between low cost, high efficiency, and small footprint was imperative to the overall goal of developing a custom PCB for GPMS.

Given the Raspberry Pi 4B's requirements, the TPS61022 presents itself strongly to be the optimal and best fit. It provides sufficient current capacity of 3A at the required 5V output, with a high efficiency that minimizes power loss. Its lower BOM cost and smaller area make it more suitable for compact, cost-sensitive applications like GPMS.

The stand-out features regarding the TPS61022 regulator are listed in the table below. [26]

Stand Out Features	Metrics
Input Voltage Range	0.5V to 5.5V
Minimum Input Voltage For Start-Up	1.8V
Output Voltage Setting Range	2.2V to 5.5V
MOSFETs Count	Two 12-mΩ (LS) / 18-mΩ (HS)
Valley Switching Current Limit	8A
Efficiency	94.7% efficiency at V_{in} = 3.6V, V_{out} = **5V and I_{out} = 3A**
Switching Frequency	1-MHz switching frequency when V_{in} > 1.5 V and 0.6-MHz switching frequency when V_{in} < 1 V
Reference Voltage Accuracy	±2.5% reference voltage accuracy over –40°C to +125°C
Pin-Selectable Mode	Pin-selectable auto PFM operation mode or forced PWM operation mode at light load
Pass-Through Mode	Pass-through mode when V_{in} > V_{out}
Shutdown	True disconnection between input and output during shutdown
Protections	Output overvoltage, thermal shutdown protections, and Output short-circuit protection
Pin Package	2-mm × 2-mm VQFN 7-pin package

Table 3.2.9.5 *Selected Regulator Metrics*

Given that the TPS61022 is an efficient power supply solution well-suited for powering devices like the Raspberry Pi 4B, the authors definitively selected this regulator to incorporate into the custom PCB. Its features begin with a broad input voltage range of 0.5 V to 5.5 V and a high switch current limit with a minimum of 6.5A.

Its capability to operate from various power sources, including batteries, makes it versatile for the Raspberry Pi 4B that may require backup power. The IC's operation at a 1-MHz switching frequency (adjustable down to 0.6 MHz for lower input voltages) ensures efficient power conversion, and its low quiescent current (26 μA) in light load conditions helps conserve energy, crucial for the battery-powered PCB. [26]

Additionally, its small footprint and minimal external component requirements fit well with the compact design of the custom PCB. With built-in protections such as output overvoltage, short-circuit, and thermal shutdown, the TPS61022 provides a reliable and efficient power solution for the Raspberry Pi 4B, enhancing its flexibility for GPMS. [26]

3.2.10 PSU Comparison

Given that the custom PCB must have a reliable form of a Power Supply Unit of its own (PSU) it will utilize batteries of a select Chemistry to supply power to the Raspberry Pi 4B and surrounding peripherals. After narrowing down the options, the authors decided to carefully put three battery types under specific consideration. Each of their specifications are explicitly listed in the table below. [27]

Chemistry	Description	Fully Charged Voltage (V)	Fully Depleted Voltage (V)	Capacity (mAh)
Alkaline	Non-rechargeable AA battery	1.65	1.4	1800 - 2700
NiMH	Rechargeable AA battery	1.45	1.2	1700 - 2000
Lithium	Non-rechargeable AA battery	3.6	3	1400 - 2500

Table 3.2.10.1 *Battery Specifications Comparison*

A comprehensive comparison of the three potential battery types are as follows with an overview of the respective pros and cons. [18]

Chemistry	Pros	Cons
Alkaline	• Lower cost per unit: Generally the cheapest option. • Readily available: Found in most stores. • No maintenance required: Simply use and dispose. • Longer shelf life (up to 10 years): Good for devices used infrequently.	• Disposable: Not environmentally friendly. • Lower capacity (mAh): May need replacing more often. • Inconsistent voltage: Voltage drops as the battery drains, which can affect MCU performance.
NiMH	• Longer shelf life (up to 10-12 years): Great for long-term storage. • Higher voltage (3.6V): Can provide more power or longer runtime for MCUs designed for that voltage. • Consistent voltage: Similar to NiMH, it offers a stable voltage output.	• Higher initial cost: More expensive than Alkaline upfront. • Requires a charger: Adds another device to manage. • Shorter shelf life (up to 5 years): Needs to be recharged periodically even if not used. • Memory effect (if not properly maintained): Can reduce capacity if not fully discharged before recharge.
Lithium	• Longer shelf life (up to 10-12 years): Great for long-term storage. • Higher voltage (3.6V): Can provide more	• Highest cost per unit: Most expensive option among the three. • Disposal needs care: Requires specific recycling

	power or longer runtime for MCUs designed for that voltage. ● Consistent voltage: Similar to NiMH, it offers a stable voltage output.	due to environmental concerns. ● Not universally compatible: Voltage might not be suitable for all MCUs

Table 3.2.10.2 *Overall Battery Comparison*

Non-rechargeable batteries, such as Alkaline and Lithium, have a lower initial purchase cost than rechargeable ones, making them more affordable for immediate use. They also exhibit a meager self-discharge rate, ensuring they retain their charge for a long time when unused, contributing to a long shelf life.

Alkaline batteries are readily available in many stores and outlets, which makes them easily accessible for the authors to attain. Regarding pricing, those of type NiMH and Lithium run at much higher costs. Therefore, fully charged Alkaline batteries, which would supply 1.65V per battery when “fresh off the shelf”, are the clear choice regarding functionality and cost, given that they are moderately priced. [28] [29] [30]

The authors explored 3 common brands of alkaline batteries which are Duracell, Energizer, Rayovac. Duracell is a well-known brand known for its long-lasting performance and dependable batteries. They offer a variety of alkaline battery types and sizes, including AA, AAA, C and D. Duracell batteries are often a good choice for devices with high power consumption or for extended use. Energizer is another popular brand of alkaline batteries. They also offer a wide range of battery types and sizes, and their batteries are known for their reliability and durability. Energizer batteries are a well all-around choice for a variety of devices. Rayovac is a third major brand of alkaline batteries. They offer a variety of battery types and sizes at competitive prices. Rayovac batteries may be a good option for devices with lower power consumption or for a simple occasional use.

Brand	Description	Strengths	Weaknesses
Duracell	A well-established brand known for durability and long-lasting performance	Reliable and long-lasting power - Offers a variety of battery types and sizes - Often a good choice for high drain devices	Can be more expensive than some other brands

Energizer	A popular brand known for its dependable and long-lasting batteries	Reliable and durable - Wide variety of battery types and sizes * Good all-around choice for many devices	Might not be the most cost-effective option for low drain devices
Rayovac	A value-oriented brand offering competitive prices	Affordable option - Available in various battery types and sizes	Performance may not match top brands in high drain devices. Might not be as durable as some other options

Table 3.2.10.3 *Alkaline Brand Battery Comparison*

Duracell utilizes a technology called "POWER BOOST" ingredients in their batteries, which claims to enhance performance in high drain devices. They also boast a 12-year storage guarantee on their Coppertop line, ensuring they'll be ready when needed.

Energizer's MAX line of batteries is known for extended lifespans. They also claim their batteries can hold power for up to 10 years in storage, and some Energizer batteries are designed to prevent leakage from damaged devices for up to 2 years.

Rayovac offers a line of Fusion Alkaline batteries designed for high drain devices. However, their overall brand might not be seen as performance-oriented compared to Duracell or Energizer. Rayovac is a solid option for those looking for a balance between affordability and decent performance for everyday use.

In summary, the authors will use the Energizer brand given that they currently obtain a fresh pack of double AA Alkaline Energizer batteries. The fact that they already obtain them will ultimately save on the overall cost of the GPMS BOM. Energizer batteries, known for their reliability and consistent performance, offer various options such as Alkaline type, suitable for a custom PCB in order to power devices like the Raspberry Pi 4B requiring sustained power. Energizer's compatibility with common AA size and seamless integration with a custom PCB, makes them extremely convenient. These batteries often deliver excellent longevity and power output, while their cost-effectiveness, considering factors like usage frequency and longevity throughout the entire development process of GPMS. [31] [32] [33]

3.3 Software Comparison and Selection

3.3.1 Operating System Comparison

Every computer requires an operating system (OS) to function effectively. Given that the computer the authors were working with was brand new, selecting the appropriate OS became a crucial decision. Their familiarity with Linux prompted them to investigate whether there were compatible drivers for an AMD GPU, as Linux was their preferred choice.

Below compares the OS options for GPMS. [34]

Feature	**Linux**	**Windows**	**Unix**
AMD Processor Compatibility	Extensive support for AMD processors	Full support for AMD processors	Varies depending on Unix variant
Cost	Mostly free and open source	Paid licenses	Paid licenses
Ease of use	Requires command line knowledge	Requires command line knowledge	Requires command line knowledge
Customizability	Highly customizable	Limited customization	Highly customizability
Security	Generally considered secure	Secure with regular updates	Generally considered secure
Stability	Highly stable	Stable with regular updates	Highly stable
Stable Diffusion Compatibility	Highly compatible	Compatible with some setup	Varies depending on Unix variant
Required Dependencies	Easy to install dependencies	Requires manual installation of dependencies	Varies depending on Unix variant
Community Support for Stable Diffusion	Large community and resources for running Stable Diffusion	Smaller compared to linux	Limited community support

Community Support in general	Large and active community	Large community and Microsoft support	Smaller, specialized communities
Scalability	Highly scalable	Scalable with some limitations	Highly scalable
Cloud integrations	Well suited for cloud environments	Integrates with Azure cloud environments	Varies depending on Unix variant
Typical use cases	Web servers, databases, cloud computing	Active directory, exchange server, .NET applications	High performance computing, mission critical applications

Table 3.3.1 *Operating System Comparison*

After carefully analyzing the comparison chart, it becomes evident that Unix ranks last among the three operating system options. Its compatibility and support vary significantly depending on the specific Unix variant, making it a less suitable choice for *GMPS's* purposes. The decision ultimately comes down to choosing between Windows and Linux.

While both operating systems have their strengths, Linux stands out in several key areas. Firstly, Linux is mostly free and open-source, providing a cost-effective solution. Secondly, it offers extensive compatibility with AMD processors and has a large community and resources dedicated to running Stable Diffusion. In contrast, Windows requires paid licenses and has a smaller community compared to Linux when it comes to Stable Diffusion support. Additionally, Linux's open-source nature allows for greater customization and flexibility. Considering the author's existing familiarity with Linux and the compelling features it offers, the authors have confidently selected Linux as the operating system of choice for the project. [34]

3.3.2 Linux Distribution and Server Options

With confirmation of Linux, the next step was to determine which distribution, or distro, would best suit the project's needs. Linux distributions encompass variations of the Linux kernel bundled with additional software, serving different purposes such as general use, server hosting, software development, and IoT applications.

Given their intention to remotely access their computer to interface with Stable Diffusion, which essentially meant hosting a server, the authors focused on distros optimized for server hosting. Among the plethora of options — including

Ubuntu, CentOS, Debian, and Fedora — they narrowed down their choices to both Ubuntu and CentOS, recognized for their suitability in this context. Below is a table comparing both. [35] [36]

Criterion	Ubuntu	CentOS
General Overview	User-friendly interface, designed for both general use and servers, with strong support for AI applications.	Enterprise-grade stability and security, derived from RHEL, suitable for server hosting and secure environments.
Intended Use	Optimized for server hosting, software development, and AI applications like Stable Diffusion.	Primarily used for server hosting with an emphasis on stability and security.
Driver Support	Compatible with AMD GPU, supports a wide range of hardware.	Compatible with AMD GPU, prioritizes stability and may have more limited hardware support compared to Ubuntu.
Update Frequency	Regular updates, including a Long-Term Support (LTS) version, though with slight delays for LTS.	Provides LTS options with regular updates, but adopts new features at a slower pace than Ubuntu.
Software Availability	Extensive range of available software, advantageous for diverse project requirements.	More limited software range, focused on stability and security.
Resource Consu-mption	Slightly higher resource consumption, which could impact performance.	Generally lower resource consumption, optimizing stability and performance for server tasks.
Community Support	Robust community support, beneficial for timely assistance and compatibility issues.	Less extensive than Ubuntu, but still offers significant support through its association with RHEL.

Table 3.3.2 *Linux & Server Comparison*

After carefully considering the various Linux distributions available, the authors have concluded that Ubuntu is the most suitable choice for the GPMS project. This decision was reached after a thorough comparison between Ubuntu and CentOS, two popular distributions known for their server hosting capabilities and compatibility with AI applications like Stable Diffusion.

Ubuntu's user-friendly interface, coupled with its strong support for AI applications, makes it an ideal candidate for the project's needs. Its optimization for server hosting, software development, and AI applications aligns perfectly with the goals of GPMS. Furthermore, Ubuntu's compatibility with AMD GPUs and its wide range of supported hardware ensures that the project can be implemented without facing significant driver-related issues.

One of the key factors that led to the selection of Ubuntu is its regular update cycle, which includes Long-Term Support (LTS) versions. Although there may be slight delays in the release of LTS updates, this approach ensures that the project can benefit from the latest features and security patches while maintaining a stable and reliable environment. The extensive range of software available for Ubuntu is another significant advantage, as it allows the project to adapt to diverse requirements and incorporate additional functionality as needed. While CentOS is known for its enterprise-grade stability and security, its more limited software range and slower adoption of new features compared to Ubuntu make it less suitable for the GPMS project. The authors believe that the benefits of Ubuntu's robust community support outweigh the slightly higher resource consumption, as timely assistance and compatibility issues can be more easily addressed with the help of the large and active Ubuntu community.

The authors also recognize the importance of staying current with advancements in the field of AI and server hosting. Ubuntu's more frequent updates and feature additions align with this goal, ensuring that the GPMS project can continue to evolve and incorporate the latest developments in the field.

Therefore, after a thorough analysis of the available options, the authors have determined that Ubuntu is the optimal Linux distribution for the GPMS project. Its combination of user-friendliness, strong support for AI applications, regular updates, extensive software availability, and robust community support make it the ideal choice for hosting the Stable Diffusion server and meeting the project's diverse requirements. By selecting Ubuntu, the authors are confident that they can create a stable, secure, and feature-rich environment that will support the successful implementation and ongoing development of the GPMS project.

3.3.3 Local vs Cloud Server

Considering a Raspberry Pi is unable to run Stable Diffusion, the authors need to look into server options. The purpose is to be able to send the data from

Raspberry Pi to the server and back. There are two main options: a local server or a cloud server. The table below shows the comparison between the two. [37]

Features	Local	Cloud
Setup	Requires configuration of server software, wireless network, and additional setup for remote access tools (VPN and SSH)	Requires configuration of server software and remote access, but the wireless network is not applicable
Wireless connectivity	Requires setting up a wireless network for local communication, Wireless range limited by local network's coverage	Not applicable, the server is accessed over the internet
Latency	Low latency as communication occurs within the local wireless	Latency depends on internet connection speed and server location
Scalability	Limited scalability based on the AMD computer's hardware capabilities	Easily scalable
Maintenance	Requires physical access to AMD server for maintenance and updates	Requires remote server management or provider maintains
Accessibility	Accessible within the range of local wireless network	Accessible from the internet anywhere
Cost	Initial cost of setting up AMD computer as a server	Recurring costs based on provider

Table 3.3.3 *Local vs Cloud Server*

Considering the fact that the AMD computer is on the school network, the authors believed it would be a lot of paperwork to expose a port to be able to use the Cloud Server. The only option from there is to work locally accessing the server by sshing into it.

After careful consideration, the authors determined that setting up a local server would be the most suitable approach for their specific use case. The AMD computer, which will act as the server, is connected to the school network. Utilizing a cloud server would require exposing a port on the school network to enable communication between the Raspberry Pi and the server. However, this process would likely involve significant paperwork and administrative overhead due to the school's network policies and security considerations.

Given these constraints, the authors opted to work with a local server setup. The Raspberry Pi will communicate with the AMD computer over the local wireless network. This local server approach simplifies the setup process, reduces latency, and eliminates the need for exposing ports on the school network.

Although a local server setup may have limitations in terms of scalability and accessibility compared to a cloud server, it provides a pragmatic solution given the constraints of the school network environment. The authors can optimize the local server's performance by configuring the AMD computer's hardware and software stack to efficiently handle the image generation workload.

3.3.3 Front-End Framework

Developing a tablet application requires careful consideration of various factors, such as compatibility with the target platform (Raspberry Pi), ease of server communication, and seamless OpenCV integration. The team prioritizes frameworks that offer straightforward OpenCV implementation, as it is crucial for the core functionality of the project to work without excessive additional effort. Other important capabilities include multi-touch support, performance, and ease of use. While the team is proficient in C++ and Python, the choice of programming language is not a limiting factor.

To assist in the decision-making process, a comprehensive chart comparing various frameworks and libraries is presented below: [38][19]

Library/ Framework	**Programming Language**	**OpenCV Integration**	**Multi-touch Support**	**Performance**	**Ease of Use**
PyQt / PySide	Python	Easy (Python bindings available)	Yes	Good	Moderate
Kivy	Python	Easy (Python bindings available)	Yes	High (OpenGL-based)	Easy to moderate
Electron	JavaScript (HTML/CSS)	Moderate (requires additional libraries)	Yes (with additional libraries)	Good	Easy (for web developers)

Flutter	Dart	Moderate (requires additional libraries)	Yes	High (natively compiled)	Moderate
Qt	C++	Easy (built-in support)	Yes	High	Moderate to complex

Table 3.3.4 *Library/Framework Comparison*

Qt stands out from the other frameworks due to its built-in support for OpenCV. This native integration makes it highly convenient to incorporate OpenCV functionality into Qt applications without the need for additional bindings or libraries. Qt's extensive documentation and resources further facilitate the implementation process.

In the context of OpenCV integration, bindings refer to the software layer that enables communication between the programming language and the OpenCV library. Bindings allow developers to access OpenCV functions and data structures from within their chosen programming language. For example, PyQt/PySide and Kivy benefit from the availability of Python bindings for OpenCV (OpenCV-Python), which makes it easy to utilize OpenCV in Python-based applications. While Qt and Kivy share many similarities, they differ in their programming languages. Qt is built on C++, known for its performance and low-level control, while Kivy is based on Python, renowned for its simplicity and rapid development capabilities. Although Kivy excels in creating visually stunning and interactive user interfaces, Qt has a larger and more vibrant community. This extensive community support is highly advantageous, providing access help if issues arise.

After careful consideration, the team has decided to use Qt as their UI framework. The combination of Qt's native OpenCV support and extensive community resources makes it the most suitable choice for their project. By leveraging Qt's strengths, the team can efficiently develop a robust and feature-rich application while benefiting from the framework's performance, cross-platform capabilities, and the vast ecosystem of tools and libraries available within the Qt community.

Chapter 4 Standards and Design Constraints

4.1 Standards

4.1.1 HDMI

High Definition Multimedia Interface (HDMI) is a standard developed by a group of electronics manufacturers designed to facilitate high-bandwidth links among digital devices. Properly configured, HDMI supports 1080p high-definition video and up to eight channels of uncompressed audio, suitable for a 7.1 surround sound setup. HDMI seamlessly simplifies the connection process by minimizing the amount of cables needed and can also decrease the number of remote controls necessary for display use. In order for all users to fully utilize all HDMI has to offer, all components in their desired system must be compatible with HDMI. [39]

A HDMI port is one of the features the Epson projector offers. This port is used to receive and transmit audio and video signals from a variety of sources. HDMI is used to send High-Definition signals over a single cable. It can be used to transmit audio and video over a cable. The benefit of HDMI is that it can transfer a high bandwidth of data via a single cable. [39]

A common misconception about HDMI and digital signals is that they are both inherently superior to analog signals due to the absence of analog-to-digital conversion, which keeps the signal in a pure and unhindered state. However, HDMI signal transmission involves encoding through Transition Minimized Differential Signaling (TMDS) to maintain consistent signal integrity over the cable length. TMDS encoding reduces signal degradation by minimizing transitions between binary states and using a twisted pair cable setup where one cable carries the signal and the other an inverse copy. As a result, this allows the receiving device to decode and compensate for any loss. [39]

Furthermore, HDMI incorporates High-Bandwidth Digital Copy Protection (HDCP) to prevent piracy. HDCP is an authentication protocol between devices, using a “handshake” process to establish a secure connection. The “handshake” ensures that only authorized devices can decode the transmitted data, with continuous checks to maintain security. The Federal Communications Commission (FCC) mandates content protection in the U.S., requiring HDMI-compatible devices to support HDCP. [39]

One common issue that users run into is a “handshake” failure where the display system does not have HDCP support. Another issue that users run into is the limit to how long an HDMI cable can be. The HDMI standard requires a minimum length of 32 feet of cable; however, it is common for many users, including the authors, to use an HDMI cable of much less length for a variety of applications. [39]

4.1.2 Transmission Control Protocol (TCP)

As part of the authors’ Senior Design project, we're integrating Transmission Control Protocol (TCP) to establish a reliable client-server connection. This connection is crucial for relaying packets from the user device to the server, where the generative AI pipeline processes the data and then back to the projector. The decision to use TCP stems from its unparalleled reliability and orderliness in data transmission, which are critical for the seamless operation of the project.

TCP is extremely fundamental to Internet communications, ensuring that data packets between networked devices are delivered accurately and also in order. In the authors' project, TCP's role is indispensable. It facilitates a dependable conduit for the data exchanges between the server's AI pipeline and the user's device.

Several features of TCP are particularly beneficial for the success of GPMS. The protocol's error-checking capabilities ensure that any data corrupted during transmission is identified and retransmitted. This guarantees the integrity of the data received by the server and, subsequently, by the projector, ensuring that the AI-generated content is displayed as intended.

Moreover, TCP's sequence control is vital for maintaining the order of data packets. This is crucial for the authors because any mis-sequence in the AI-generated data could lead to incorrect or nonsensical visual outputs. Additionally, TCP's flow control mechanism helps prevent network congestion, which is essential for maintaining a smooth and responsive interaction between the user device and the server, to cut down on wait time during generation.

Diving deeper into the technical aspects of the Transmission Control Protocol (TCP) illuminates how its sophisticated mechanisms are instrumental for the project. One of TCP's standout features is its use of a three-way handshake to establish a connection. This process involves an exchange of SYN (synchronize) and ACK (acknowledge) messages between the client and server. This handshake ensures that both the sender and receiver are ready for data transmission, establishing a reliable channel. For the authors' project, this means they must create a stable pathway for data to flow between the user device and their server, which is extremely crucial for processing in the generative AI pipeline.

Furthermore, TCP's congestion control algorithms are a cornerstone of its ability to manage data flow efficiently across networks. These algorithms, such as Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery, dynamically adjust the rate of data transmission based on the network's capacity to handle traffic. This adaptability prevents packet loss and minimizes latency, which is paramount for GPMS.

The authors' generative AI pipeline relies on the timely and accurate reception of data to generate and project visuals without delay or disruption. By leveraging TCP's congestion control, the authors can ensure that the system remains responsive and efficient, even under varying network conditions, providing a seamless experience for all users.

Implementing TCP in the authors' architecture requires careful consideration. The authors need to establish robust TCP connections that can handle the intensive data exchange required by their AI applications. This includes managing TCP sessions efficiently to ensure a stable and continuous flow of information.

Despite TCP's advantages, the authors anticipate challenges, particularly in minimizing latency and maintaining real-time responsiveness. These are critical aspects that they are addressing through system design optimizations and by leveraging TCP's capabilities to their fullest. The goal is to ensure that the authors' project delivers a seamless and engaging experience powered by a reliable and efficient backend communication system.

4.1.3 Secure Shell (SSH)

As part of the authors' Senior Design project, integrating Secure Shell (SSH) is pivotal for establishing a secure connection between the client and the server. This security is essential for ensuring that the data transmitted in the project's generative AI pipeline is protected against unauthorized access or disclosure, primarily due to the nature of the machine being on UCF's network. The choice to utilize SSH was driven by its robust encryption standards, which safeguard data integrity and confidentiality during transmission—essential conditions for the security aspects of the project.

SSH is a protocol widely recognized for its strong security measures and is used predominantly for secure network services over an unsecured network. In the authors' project, the role of SSH is critical as it provides a secure channel over an insecure network, ensuring that all data sent and received remains confidential and unaltered.

Several features of SSH make it an excellent choice for securing the authors' project communications. One significant capability of SSH is its comprehensive encryption framework, which prevents sensitive data from being exposed or tampered with during transit. This encryption is vital for maintaining the privacy of the data exchanged between the server's AI pipeline and the user's device, ensuring that all communications are shielded from potential eavesdropping or interception. [40]

Moreover, SSH offers more than just data encryption; it also provides strong authentication mechanisms. This includes the use of public key authentication, which ensures that the connection is made with the intended server and not an imposter. For the authors, this means an added layer of security, guaranteeing that only authorized devices can connect and interact with the server, an essential factor in preventing unauthorized access.

Delving into the technical specifics of Secure Shell (SSH) reveals how its features are tailored to meet the demands of secure transmission. One key component of SSH is its ability to multiplex multiple virtual channels over a single network connection. This capability allows the transmission of various types of data (such as commands, files, and device forwarding information) independently but within the same secure connection. For the authors' project, this means they can manage multiple data streams efficiently without compromising security, which is important for operating their AI pipeline securely and effectively. [40]

Additionally, SSH includes mechanisms to prevent several types of cryptographic attacks, such as man-in-the-middle attacks, which are critical for the project's security stance. It uses strict host key checking and dynamic data exchange encryption to ensure that each session is securely isolated and that attackers cannot reuse the keys. [40]

The implementation of SSH in the authors' project architecture necessitates particular setup and configuration. However, these critical aspects are later discussed in chapter 7.2.4

In conclusion, the use of SSH is integral to the authors' senior design project, providing a secure communication channel that is crucial for the integrity and confidentiality of data in their generative AI pipeline. By effectively leveraging SSH's security features, the authors aim to deliver a project that not only functions efficiently but also adheres to the highest standards of data security, ensuring a secure and reliable user experience. [40]

4.1.4 Battery

Two main organizations set the standards for alkaline batteries: the International Electrotechnical Commission (IEC) and the American National Standards Institute (ANSI). These standards ensure consistent sizing, performance, and labeling across different brands and manufacturers.

IEC Standards is the first main organization. The IEC designation system uses a letter prefix followed directly by a number. The letter prefix indicates the general chemistry of the battery. For example, 'L' stands for Alkaline (the battery type that GPMS will utilize). Then, the number denotes the size of the battery. For example, AA batteries are designated LR6 according to the IEC system.

ANSI Standards is the second main organization. The ANSI designation system uses a two-letter code to denote the size of the battery. For example, some of the more common examples include AA, AAA, C, D, and 9V.

Both IEC and ANSI standards also specify performance metrics for Alkaline batteries, such as voltage, capacity, and shelf life. These standards help ensure that consumers can purchase batteries, including those of Alkaline type, with predictable performance characteristics.

Beyond the basic sizing and labeling requirements, battery standards also delve into a couple characteristics regarding safety, performance, and durability.

Firstly, regarding safety, standards address venting mechanisms to prevent explosions due to gas buildup during use or misuse. They also regulate mercury and other potentially harmful materials, ensuring safe battery disposal and environmental protection.

Secondly, regarding performance, standards define minimum performance benchmarks for capacity (mAh - milliampere-hours) and shelf life which is the storage time before significant power loss. This allows all consumers to compare the different batteries from different brands and choose the one that best suits their specific needs.

Lastly, regarding durability, standards specify testing procedures to simulate real-world usage and ensure batteries can withstand certain levels of shock, vibration, and extreme temperatures. This helps guarantee reliable performance in various applications.

In conclusion, the standards set by the IEC and ANSI play a crucial role in ensuring uniformity and quality in batteries. By establishing consistent sizing, performance metrics, and labeling protocols, these organizations enable consumers to make informed choices when purchasing batteries. Moreover, the standards encompass safety measures, performance benchmarks, and durability testing, further enhancing the reliability and suitability of Alkaline batteries for GPMS. Overall, these standards uphold consumer trust and environmental responsibility while simultaneously promoting innovation and advancement in battery technology. [41]

4.1.5 Projector

Projectors come in a variety of shapes and sizes, but they all adhere to a few key standards. To begin with, it is necessary to discuss resolutions. Projector resolution refers to the number of pixels that can be displayed on the screen. Common resolutions include SVGA (800x600), XGA (1024x768), WXGA (1280x768), and Full HD (1920x1080). The relationship between resolution and detail is directly proportional. Higher resolutions provide sharper and more detailed images, and lower resolutions provide less sharp and less detailed images.

Secondly, aspect ratio is crucial. The aspect ratio is the width of the projected image compared to its height. Common aspect ratios include 4:3 (standard definition) and 16:9 (widescreen). The aspect ratio of the projector should match the aspect ratio of the content users are projecting. Thirdly, lumens encompass a critical detail to the overall standard regarding projectors. Lumens is a unit that measures the brightness of a projector. Projectors with higher lumens ratings are better suited for use in brightly lit rooms. Given just a few of the key projector standards, users can first understand these standards and then choose a projector that best meets their needs for brightness, resolution, and compatibility. Beyond these few standards, projector standards encompass more than just resolution, aspect ratio, and lumens. They also revolve around connectivity, lamp life, throw ratio, and color.

Firstly, there exist standards to define supported connection options like HDMI, VGA, and USB for connecting the projector to your devices. Secondly, lamp life: Lamp life refers to the average number of hours the projector lamp can last before needing replacement. Knowing the lamp's life helps estimate ongoing costs associated with projector usage. Thirdly, the throw ratio determines the distance required between the projector and the screen to achieve the desired image size. Understanding throw ratio is crucial for setting up the projector in a room of specific dimensions. Furthermore, projector standards also encompass color gamut, which refers to the range of colors a projector can display. This is particularly important for applications, such as GPMS, that facilitate photo projections/presentations.

Projectors adhere to several key standards to ensure optimal performance and compatibility. Resolution, which determines the level of detail, comes in various common formats. The aspect ratio, representing the width-to-height ratio of the projected image, should match the content being displayed. Lumens measure brightness, with higher ratings suitable for brighter environments. Beyond these standards, connectivity options such as HDMI, lamp life indicating replacement frequency, throw ratio for optimal placement, and color gamut for accurate color representation are also essential considerations. By understanding and evaluating these standards, users can select projectors that best suit their specific needs and preferences. [42]

4.1.6 Camera

There are two main standards for cameras. Measurement Standards define how various aspects of camera performance are measured. They are set by organizations like the International Organization for Standardization (ISO) and ensure consistency in how camera manufacturers report things like resolution, noise, and sensitivity. Environmental Standards define how rugged a camera is in terms of dust, water, and impact resistance. They are important for understanding how well a camera will withstand different environmental conditions. After understanding these standards, the authors were more equipped to compare cameras and choose the one that best met their needs.

The ISO 12233 standard defines how a camera's resolution, often megapixels (MP), is measured. It specifies using a chart with specific patterns to test the camera's ability to capture fine details. It is important to note that a higher MP does not always equate to better image quality, but it does indeed allow larger prints or cropping without losing detail.

The ISO 15739 standard defines image noise which refers to unwanted grain or speckles that appear in photos, especially and oftentimes in low-light situations. This standard defines how to measure noise levels, allowing comparison between different types of cameras. In fact, lower noise levels indicate a cleaner image.

The ISO 12232 refers to the ISO Speed Rating. This standard defines how a camera's ISO speed, which determines its sensitivity to light, is measured and reported. Higher ISO allows for capturing images in darker environments but can also introduce more noise. This standard ensures consistent reporting across camera brands.

The ISO 14524 standard, the tone curve, defines how a camera translates light captured by the sensor into brightness levels in the image file. The ISO 17850, geometric distortion, measures how straight lines appear curved in the final image due to lens imperfections. The ISO 18844, image flare, defines how a camera handles unwanted light sources appearing as streaks or halos in the image.

The IK Rating, referring to impact resistance, is an international standard (IEC 62262) that uses a two-digit code to rate a camera's resistance to physical impacts from accidental drops or bumps. Higher numbers indicate greater impact resistance. The IP rating, referring to Ingress Protection, IP ratings (IEC 60529) define a camera's resistance to dust and water ingress. The first digit indicates dust protection, and the second digit represents water resistance. For example, IPX4 signifies protection against splashing water. The NEMA Standards (designated to North America) define enclosure ratings for electrical equipment, including cameras. Similar to IP ratings, they use a code to indicate dust and water resistance levels.

In conclusion, camera standards act as a common language for understanding a camera's capabilities. Measurement standards (ISO) ensure clear communication regarding a camera's resolution, low-light performance, and other key aspects of image quality. Environmental standards (IK rating and NEMA) inform one about the camera's durability in harsh conditions. By understanding these standards, the authors were well-informed to select a camera that best suits their needs. [43]

4.2 Constraints
4.2.1 Economic

The economic constraints of the GPMS project primarily revolve around the cost of the hardware components and the available budget. The total estimated cost for the project is approximately $3,544.37, as detailed in Section 10.1. However, thanks to component donations from friends and family, the out-of-pocket expenses for the authors are reduced to roughly $299.93.

Despite these donations, the authors recognize that additional funding could have enhanced the project's capabilities. For instance, if the budget allowed, the authors would have opted for an NVIDIA GPU instead of the AMD GPU currently being used. NVIDIA GPUs are widely known for their superior performance in deep learning and AI applications in comparison to AMD GPUs, which could have certainly improved the image generation quality and processing speed of the GPMS.

Furthermore, a higher budget would have enabled the authors to acquire a high-end projector capable of projecting images at greater distances with increased visibility. Such a projector would have expanded the potential applications of the GPMS, allowing for larger-scale projections in various settings.

While the current budget and donated components have allowed the authors to develop a functional prototype, they acknowledge that additional financial resources could have elevated the project's performance and versatility. Careful financial planning and management will be crucial to successfully completing the GPMS project within the given economic constraints while striving to maximize its capabilities.

4.2.2 Time

The GPMS project is subject to the time constraints imposed by the Senior Design joint course schedule, which presents several challenges for the project team. The entirety of the GPMS project is divided into two courses/phases: Senior Design 1 (EEL4914) and Senior Design 2 (EEL 4915L), each with its own set of deadlines and deliverables.

Senior Design 1 focuses on completing all administrative tasks, such as project planning, research, and documentation. The authors have set a series of deadlines (detailed in Chapter 10.2.1) to ensure steady progress throughout this phase. Key milestones include completing the 45-page document by March 10, 2024, completing the user interface by March 20, 2024, and submitting the 90-page document by April 23, 2024. These deadlines are crucial for laying the foundation of the project and ensuring that all necessary research and planning are completed before moving on to the implementation phase.

However, the team faces additional challenges that may impact their ability to meet these deadlines. Two of the team members have limited experience with C++ programming, which is a critical skill required for developing the software components of the GPMS project. This knowledge gap may slow down the coding process, as these team members will need to spend extra time learning and familiarizing themselves with the language and its best practices. To mitigate this challenge, the team will need to allocate additional time for training and peer support, ensuring that all members are equipped with the necessary skills to contribute effectively to the project.

Moreover, the team members are simultaneously enrolled in other classes, each with its own set of assignments, projects, and exams. Balancing the workload of multiple courses can be a significant challenge, especially when deadlines from different classes align. The final paper submission for the GPMS project, in particular, coincides with the finals week for other courses, placing additional pressure on the team members to manage their time effectively and prioritize their tasks.

Senior Design 2 is dedicated to prototyping, testing, and debugging the entire project. The authors plan to test the PCB by May 31, 2024, and complete the prototype and testing of the entire project by August 20, 2024. The final demonstration is scheduled for December 1, 2024. These milestones require significant effort and coordination from the team, as they involve integrating various hardware and software components, conducting rigorous testing, and addressing any issues that may arise during the process.

To meet these deadlines and effectively manage time constraints, the authors will need to adhere to a strict schedule, regularly communicate progress, and promptly address any issues that may arise. Effective time management and collaboration among team members will be essential to ensure the project is completed within the allotted time frame.

4.2.3 Ethics

Ethical constraints exist in the development of GPMS. Generative AI, a key component to GPMS, has transformative potential and the ethical challenges posed in image generation. On the positive side, AI image generators like DALL-E, Midjourney, and Stable Diffusion enable ease of creative thought and make digital art/design accessible to a broader audience regardless of any skill level.

This technology enables rapid iteration of images, which in turn reduces the time and cost associated with traditional methods of image content creation. From an environmental point of view, relying on AI for image generation could potentially reduce the carbon footprint associated with the physical aspects of artistic production. The use of AI image generation fosters innovation by offering new avenues for visual media. [44]

However, significant ethical concerns surround AI image generation for a number of reasons. Since AI tools, such as ChatGPT 4, often use large datasets of images without originality metrics, the risk of copyright infringement and plagiarism is a major issue. This can lead to the creation of works that infringe on the rights of original creators, posing potential legal disputes. Additionally, AI-generation images may even display biases and stereotypes based on user inputs.

This emphasizes the need for careful consideration of the social implications of these image-generation technologies. There is a prevalent and vital call for ethical guidelines and legal clarity to address the complex landscape of AI image generation. This is where the need for a balanced approach that respects both the innovative potential of AI and the rights and contributions of human creators. [44]

In the context of GPMS, the user input must prohibit inappropriate and misleading images. This is because the authors desire GPMS to be a product for all audiences. With the previous measures in place, the integrity of GPMS is ensured and intact, especially for family-friendly environments.

4.2.4 Security / Network Integration

During the research phase of the project, the team explored various methods for establishing communication between the Raspberry Pi and the AMD computer. The goal was to find an efficient and secure way to transfer data and execute computationally intensive tasks on the more powerful AMD machine. However, several constraints and challenges emerged throughout the process.

One of the primary hurdles encountered was the limited access to the AMD computer. The team collaborated with a UCF employee who provided assistance in granting access to the machine. However, the process experienced significant delays due to unforeseen issues and administrative complexities regarding UCF IT. These hindered the team's progress and raised concerns about the feasibility of the proposed communication system, specifically the hopes of exposing a port on the network to simplify API (Application Programming Interface) calls.

A major constraint that came to light was the restricted accessibility of the AMD computer. The machine could only be accessed remotely through SSH (Secure Shell) while connected to the school network. This limitation posed a significant challenge, as it required team members to be physically present on campus to establish a connection. Although this issue could be mitigated by using Cisco AnyConnect, a VPN (Virtual Private Network) solution employed in other areas to access school resources remotely, it introduced additional complexity.

The most critical issue, however, stemmed from the inability to make conventional API calls to the AMD computer. In order to facilitate seamless communication and data exchange, the team initially planned to utilize API endpoints. However, this approach required opening a specific port on the school network to listen for incoming requests. Given the strict security policies and administrative regulations in place, particularly those related to HIPAA (Health Insurance Portability and Accountability Act), the team realized that obtaining permission to open a dedicated port would be a formidable challenge.

The school's IT department and administration maintain stringent control over network access and security measures. The lengthy process of acquiring SSH permission for the project served as a clear indication of the level of scrutiny and caution exercised by the institution. Consequently, the team concluded that securing approval to open a port for API communication would be highly unlikely, considering the time-consuming nature of the SSH permission process alone.

This constraint forced the team to reevaluate their approach and explore alternative solutions for enabling communication between the Raspberry Pi and the AMD computer. The inability to rely on standard API calls necessitated the development of creative workarounds and unconventional methods to facilitate data transfer and remote execution of computationally intensive tasks.

In the following sections of the paper, the team will delve into these alternative methods and workarounds they explored to overcome the constraints associated with the machine. These solutions will provide a secure and efficient means of data transfer and remote execution while respecting the school's network policies and ensuring the integrity of the project. By thoroughly examining the constraints related to the AMD computer, the team gained a deeper understanding of the complexities involved in integrating external resources into their project.

Chapter 5 Comparison of ChatGPT with other Similar Platforms

5.1 Comparison Overview

In recent years, generative AI platforms like ChatGPT, Microsoft's Copilot, and Gemini have revolutionized how individuals and organizations approach problem-solving, creativity, and learning. These platforms, through their advanced algorithms and vast databases, have the capability to assist and guide ambitious individuals, such as the authors, in the development of complex projects.

These platforms can directly impact Senior Design Engineering capstone projects across the country, including GPMS. This chapter aims to highlight and compare each of these platforms, focusing on their limitations, pros and cons, and impact on learning experiences within the context of the authors’ experience creating GPMS.

5.1.1 ChatGPT

ChatGPT is a free AI-powered natural language processing tool that directly enables human-like interactions and offers various capabilities with the chatbot. This language model can respond to inquiries and help with tasks, including writing emails, essays, and code. [45]

ChatGPT, developed by OpenAI, is widely known by technology users for its conversational abilities, making it an excellent tool for generating ideas. That being said, it is also helpful in suggesting code optimizations and alternative solutions for problems. Creating content is also a lucrative feature regarding ChatGPT's capabilities ever since its newest paid subscription model, ChatGPT-4, was released. This newest model is capable of generating AI-generated images in seconds. [45]

In the GPMS research process, the authors pondered whether to use a basic search engine to begin the research process or turn directly to AI. Understanding the differences was paramount to taking the next steps. It was understood that ChatGPT is designed to engage users in dialogue, whereas search engines like Google index web pages to deliver requested information. Each serves a distinct purpose, making them incomparable in terms of utility. [45]

However, search engines are preferred for the latest and most accurate information. That is because most ChatGPT users opt for the basic free version of ChatGPT, known as ChatGPT-3.5, which lacks internet search capabilities, relying instead on pre-existing knowledge from its training data, which may introduce inaccuracies. [45]

ChatGPT's free version only has knowledge up to 2021, limiting its ability to answer questions about events occurring after that year. On the other hand, search engineers like Google offer the most recent information. ChatGPT-4 subscribers benefit from Bing integration, allowing the chatbot internet access. That enhances ChatGPT's functionality with web indexing while maintaining its capability to process natural language queries and generate conversational responses. [45]

Given ChatGPT's ability to explain complex topics in a simplified manner, it was the most suitable option for GPMS research. Furthermore, while ChatGPT was extremely helpful, it revealed several limitations. Users may need to rephrase questions multiple times to convey their intent clearly, and the quality of responses can vary. In fact, in an attempt to facilitate a conversation regarding GPMS, this platform tended to make assumptions, leading to misinterpretations. [45]

A significant limitation in the free version of ChatGPT is the knowledge cutoff in 2021. This gap leaves the platform needing to be made aware of more recent events or developments. Additionally, the free version of ChatGPT initially did not cite sources for its information. However, this has been partially addressed by developers for ChatGPT-4 users, given that the newer version can aid in proper citation generation when needed. [45]

ChatGPT-4 was imperative and crucial in generating concise and precise tables for the GPMS technology comparison section of this document. ChatGPT-3.5, the free version, was not capable of generating tables the way the paid version did. More specifically, tables were only generated via ChatGPT-4 using paragraphs that contained detailed explanations and comparisons of at least three technology parts.

The tables returned from ChatGPT-4 contained a complete and concise overview that quickly ran through a set of factors that pertained to the specific technology part at hand, unlike the tables that ChatGPT-3.5 returned. This comprehensive capability saved much time from manually converting pre-existing comparison content into concise tables comparing three technology parts. Instead of spending time on manual table generation, more time was redirected to performing minor adjustments and polishing the AI generated tables.

5.1.2 Copilot

Microsoft has been a key player in generative AI, which aims to produce content from user input. It has added AI capabilities to Bing and introduced an early version of the Copilot AI assistant in Windows 11, designed to boost both creativity and productivity.

Copilot, initially known as Bing Chat, is Microsoft's advanced AI chatbot integrated into its search engine and is powered by a technology more sophisticated than OpenAI's GPT-4. It was unveiled in early 2022 and later rebranded during the Microsoft Ignite event.

Copilot can assist users with various tasks, including creative writing, solving math problems, coding, and generating images. It stands out because it can provide up-to-date information from the web with links, which was extremely helpful in researching various parts of GPMS. In fact, Copilot has a feature that allows users to choose a conversation style such as "More Creative, "More Balanced," or "More Precise." These styles bring character limits. The "More Creative" and "More Precise" styles allow 4000 characters in total text input. Unfortunately, the "More Balanced" style only allows a total of 2000 characters in total of text input. [46]

Additionally, Copilot can accept both text and image inputs. This was essential to developing the ideal cover image of GPMS documentation. It is interesting to note that Microsoft's Copilot had an autocomplete feature with text inputs, which was extremely helpful in brainstorming specific adjectives to describe the desired image. A comparison between a ChatGPT image and a Copilot image resulted in utilizing the ChatGPT-generated image for simplicity. [46]

Microsoft also uses the Copilot branding for its range of AI assistants tailored to specific services, like Copilot for Windows, Copilot for Microsoft 365, and more. The tool will receive new features such as GPT-4 Turbo integration, enhanced search capabilities, and Code Interpreter. [46]

Copilot is accessible via its direct website to anyone with a Microsoft account, and its use cases are diverse, including functioning as a search engine, content creator, and learning tool from uploaded images. A Copilot app is available for mobile devices, and a premium version called Copilot Pro offers expanded features for a subscription fee. [46]

Microsoft, an early investor in OpenAI, has incorporated ChatGPT into Bing, leveraging their partnership. Microsoft's Azure is OpenAI's exclusive cloud-computing provider. Initial controversies around the AI chatbot's behavior led to the implementation of chat limits.

Despite these issues, Copilot has been rated highly for its capabilities, including access to GPT-4 and internet connectivity. Copilot solved some of the significant problems ChatGPT faces, such as including knowledge of current events via internet access and providing footnotes with links to sources from the information it pulled. [46] Additionally, Microsoft offers Bing Chat Enterprise, now called Copilot, as a secure workplace chatbot solution, initially provided at no extra cost to certain Microsoft 365 account holders, with a planned future fee.

5.1.3 Gemini

Gemini, formerly known as Google Bard, is Google's conversational AI chatbot designed to rival ChatGPT. Unlike ChatGPT, which relies on information up to 2021, Gemini extracts real-time data from the web. This in itself is a massive reason why Gemini is beneficial for research. It can code, solve math problems, assist with writing, and, starting February 2024, generate images using Google's Imagen 2 model. [29]

Originally launched as Bard and based on a lightweight version of LaMDA, Gemini has seen significant upgrades. Initially, it utilized Google's Transformer neural network architecture but later transitioned to PaLM 2 for enhanced performance. In December 2023, Google introduced an advanced version of Gemini, leveraging a fine-tuned variant of Gemini Pro for English, marking it as Google's most capable Language Learning Model. [29]

Gemini is publicly available without a waitlist. It supports multimodal search, integrating Google Lens to allow users to input images alongside text for queries. This feature enables users to receive information about images they upload, such as identifying plants or even breeds of pets. [29]

Additionally, Gemini can include images in its responses for queries that benefit from visual aids. With the introduction of image generation capabilities, users can request Gemini to create images based on detailed or minimal descriptions, showcasing Google's advancements in AI-driven creativity. [29] Gemini represents Google's ambitious foray into conversational AI, offering real-time web information retrieval, multimodal searches, and advanced image generation powered by its continually evolving large language models. [29]

So far, Gemini has been an extreme aid in translating existing technology part comparisons into tables to best visualize each given part's pros and cons. As previously discussed, Gemini aided much more substantially in generating tables derived from pre-existing paragraphs describing technology components compared to the ChatGPT-3.5 Version. This is because Gemini provided a comprehensive overview that contained sources for the information of each technology part. With regard to ChatGPT-4, Gemini equally matched the quality of the table generation with regard to accuracy and source provision.

5.1.4 Overall Comparison

The table below provides a comparison of various aspects of ChatGPT, Copilot, and Gemini. Each of their functionalities, features, strengths, weaknesses, and other relevant details are highlighted.

Aspect	ChatGPT	Copilot	Gemini
Core Functionality	Natural language processing, conversation, idea generation	AI assistant, creativity booster, productivity enhancer	Conversational AI, real-time web data extraction
Developer	OpenAI	Microsoft	Google
Internet Access	Limited (ChatGPT-3.5 lacks internet access)	Yes (provides up-to-date information with links)	Yes (extracts real-time data from the web)
Subscription Model	Free (with limited features)	Free (premium version available for a subscription	Free

		fee)	
Image Generation	Yes (Only ChatGPT-4 can generate AI-generated images)	Yes (supports both text and image inputs)	Yes (supports image generation capabilities)
Collaboration	Limited	Yes (offers Copilot for Microsoft 365)	Not specified
Multimodal Support	No	Yes (accepts both text and image inputs)	Yes (supports Google Lens for multimodal search)
Search Capabilities	Limited to pre-existing knowledge	Enhanced (access to GPT-4 and internet connectivity)	Real-time web data extraction
Integration	Standalone tool	Integrated into Bing and Microsoft ecosystem	Integrated into Google ecosystem
Table Generation	Limited (ChatGPT-3.5 lacks decent table generation capability)	Yes (Copilot Pro offers advanced features)	Yes (supports comprehensive table generation)
Knowledge Cutoff	Limited to 2021	Provides up-to-date information from the web	Real-time web data extraction
Strengths	Conversational abilities, code suggestions, idea generation	Up-to-date information, internet connectivity, code assistance	Real-time web data extraction, image generation
Weaknesses	Limited internet access, outdated knowledge	Initial controversies, chat limits, subscription	None experienced

		model	

Table 5.1.4 *ChatGPT vs Copilot vs Gemini*

The table above presents a comprehensive comparison of three leading generative AI platforms: ChatGPT, Copilot, and Gemini. Despite each of their differences, each specific platform offers unique strengths, from conversation and code suggestions to real-time web data extraction and advanced image generation.

5.2 Limitations, Pros, and Cons

ChatGPT will sometimes generate inaccurate information and has been known to struggle with highly specialized or niche topics. In order to use ChatGPT, users must create an account and then login before accessing any of the services that ChatGPT's free or paid version provide.

The main pro of ChatGPT is the fact that it is highly versatile in content creation; excellent at understanding and generating human-like text. However, its main con is that it is limited real-time data access; potential for generating incorrect information.

Microsoft Copilot is limited primarily to the Microsoft ecosystem. Furthermore it may not always understand the context or the specificity of user requests outside of its integrated applications. Microsoft Copilot requires all of its users to first create an account and then login before accessing some of the services that it provides such as image generation.

The main pro of Microsoft Copilot is its deep integration with Microsoft tools; real-time assistance within familiar applications. Despite this pro, it is important to know that the main con is that it is less versatile outside of the Microsoft ecosystem; potential for dependency on specific software.

Google Gemini, formerly known as Google Bard's, reliance on web-based information can lead to biases or inaccuracies. Furthermore, it may not provide as personalized or context-aware responses as other platforms. Gemini requires all of its users to first create an account and then login before accessing any of the services that it provides.

The main pro of Google Gemini is that it has a plethora of access to a vast range of information; innovative in generating insights based on current data. A quite noticeable drawback is that it does have potential for bias in information, which in return gives less focus on generative creativity in direct comparison to conversational understanding.

5.3 Impact on Senior Design Engineering Capstone Projects

The platforms above cover various areas that directly impacted the refinement of GPMS.

Through ChatGPT, multiple ideas for the authors' Senior Design Capstone Project were generated. Moreover, ChatGPT began the authors' brainstorming process. This widely-known platform also aided in suggestions based on similar projects and current trends.

Regarding Microsoft Copilot, the authors experienced significant benefits regarding software integration by offering real-time coding assistance and debugging tips. It also provided PSU recommendations regarding using the most optimal batteries to kickstart the custom PCB development process.

Gemini aided in conducting thorough research by pulling in the latest studies, articles, and data from across the Internet, ensuring that GPMS was grounded in up-to-date information with supporting documentation when necessary. This was especially important given that it was essential to the authors that the project was in the realm of possibilities and could realistically be completed.

Although beneficial in most cases, these platforms have the potential to unintentionally impose harmful effects. These effects include, but are not limited to, overreliance, data privacy concerns, and skill atrophy. An overreliance on the AI platforms above can hinder the development of personal critical thinking and problem-solving skills among a group. Additionally, group collaboration has the potential to decrease in the presence of AI use. Fortunately, AI only improved the consistency of group collaboration.

Using the platforms for GPMS or any other project development could raise concerns regarding sensitive data privacy. This is because when sensitive or proprietary information is involved, it raises concerns on whether authors should be inputting their project-specific information into any select platform. Regarding skill deterioration, the convenience of AI assistance could lead to a decline in one's ability to perform time-consuming tasks independently, such as manual coding or original content creation. Fortunately, the authors did not run into these select problems.

5.4 Concluding Remarks

ChatGPT, Microsoft Copilot, and Google Gemini each offer unique advantages and face distinct limitations. Their impact on Senior Design Engineering capstone projects, such as GPMS, can be profound, offering tools for innovation, efficiency, and creativity. However, it's essential for each group member to be mindful of the potential drawbacks, such as overreliance and the risk of diminishing critical thinking skills that can otherwise benefit the group.

By leveraging these platforms, each group member can enhance their learning experiences while fostering a balance between AI-assisted efficiency and the development of independent, critical Engineering skills required for the overall success of developing GPMS.

Chapter 6 Hardware Design

6.1 Custom PCB Overview

The custom PCB was designed using Autodesk Fusion 360's electronic schematic interface. A detailed strategic layout was paramount to designing the custom PCB in order to create a comprehensive and efficient schematic that all readers can easily follow.

In the first iteration of designing a schematic for the custom PCB, two battery holders will each house two Alkaline batteries and will jointly serve as the input voltage V_{in} of the PSU to the Raspberry Pi 4B. The two pairs of batteries are followed by an on/off switch to allow/not allow current to flow when switched on/off. A red LED will shine to indicate when the switch is on. When the LED is off, the switch is off. Following the switch is a regulator, more specifically, a DC/DC converter. The regulator steps up the incoming voltage to the Pin Header which will be used to connect to the Raspberry Pi 4B input Voltage pins.

A single battery holder that contains 2 Alkaline batteries in series will produce a 3.3V power supply. When a second battery holder containing 2 Alkaline batteries is put in parallel to the first battery holder, the voltage remains constant at 3.3V.

When you connect batteries in parallel, the positive terminals are connected together and the negative terminals are connected together. This setup maintains the same voltage as a single battery but increases the overall capacity of the battery bank. Voltage remains constant because the potential difference between the positive and negative terminals of each battery stays the same.

Since the two battery holders are placed in parallel, each containing two Alkaline batteries, does not change the original voltage but actually produces more current flow throughout the custom PCB, through the regulator, and into the Pi. This increased current will be effectively regulated through the power leads on the Raspberry Pi 4B.

The DC/DC converter directs/steps up the incoming voltage from the batteries to the 5V voltage level that the Raspberry Pi 4B requires and into the 3 pin header that follows the regulator. A female-to-female jumper wire will be used to connect the top most pin of the 3 pin header to the corresponding 5V input voltage pin of the Raspberry Pi 4B. The middle and bottom most pins of the 3 pin header are the ground pins and will be connected to the Pi's corresponding ground pins via female to female jumper wire.

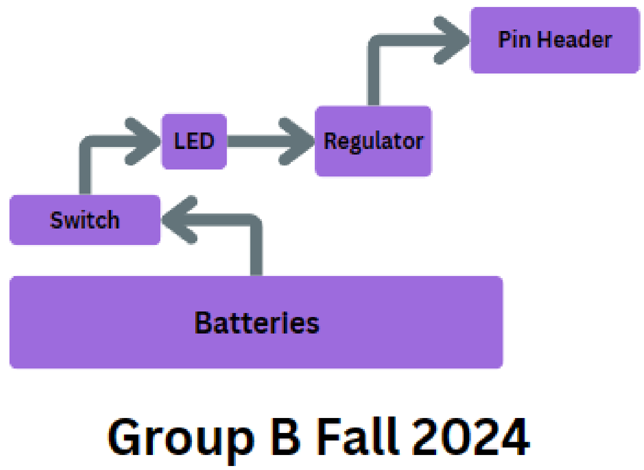


Figure 6.1 *Custom PCB Subsystem Block Diagram*

Given that The Raspberry Pi 4B requires an input voltage of 5V, pin 2 of the GPIO layout on the board will be utilized which is visible on Figure E. The middle and bottom pins of the 3-pin header will be directly connected to pin 6 and pin 20 respectively for the ground connections. The ground pins on the Raspberry Pi 4B are also visible on Figure E.

6.2 Custom PCB Schematics

The schematic of the custom PCB's first revision is listed below in Figure 6.2.1. As shown below, there are two battery holders in parallel with each other with the negative terminals of each battery holder directly connected to ground.

When the switch is turned *off*, no current will flow through the switch. When the switch is turned *on*, current will first flow through to a resistor to limit itself before reaching the red LED. When the current reaches the red LED, it will turn on to indicate it is on. Then the current will pass through a DC/DC converter to up the voltage to 5V at 3A, which is what the Raspberry Pi 4B requires. [47] The output of the DC/DC converter goes to a 3 pin header. [13]

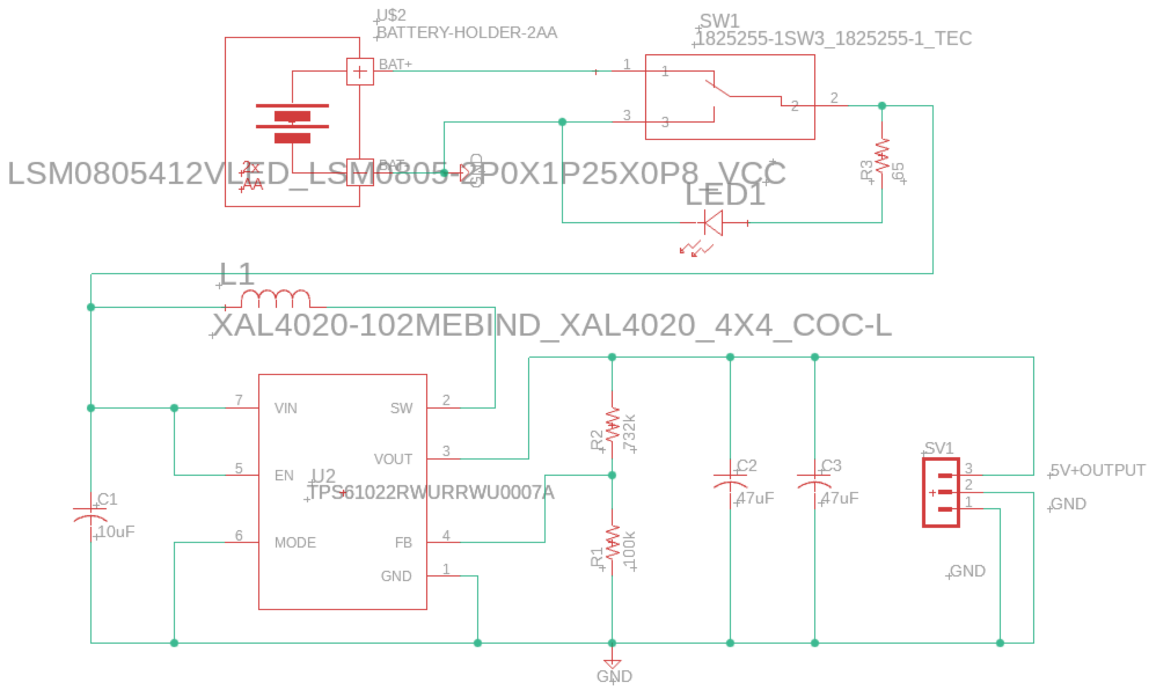


Figure 6.2.1 *Custom PCB Schematic (First Revision)*

Given the 3.3V power supply from freshly unused Alkaline batteries, a resistor can be strategically used to limit the current flowing through the red LED. Unfortunately, if the incorrect resistor value is used, the current will not be limited, and the LED will consequently burn out. Ohm's Law was used to determine the resistor value required for the LED to avoid burning out using its forward voltage and forward current.

The equation needed to calculate the correct resistance value for each LED to avoid burning is:

$$R = \frac{(V_{source} - V_{forward})}{I_{forward}}$$

Figure 6.2.2 *Ohm's Law*

$$R = \frac{(3.3V - 2.0V)}{20mA}$$

where V_{source} is the voltage of the two Alkaline batteries, $V_{forward}$ is the forward voltage of the diode, and $I_{forward}$ current of the LED.

The red LED, which indicates that the power is on/off, has a forward voltage of 2.0V and a forward current of 20mA. Using these values, the resistor value is then computed as at approximately 65Ω using Ohm's Law which is shown above in Figure 6.2.2. [48] A red LED was chosen because the custom PCB would be manufactured in green color. Therefore, the authors selected an LED that best contrasted the custom PCB. To summarize, 3.3V will not fry the LED component of the custom PCB if and only if the resistor is used to limit the current to a safe level.

The output of the DC/DC converter will flow into a 3-pin header where the middle and bottom most pins of the 3-pin header will connect to the ground pins of the Raspberry Pi 4B via female-to-female jumper wires, while the remaining pin will connect to the 5V Vcc pin on the Pi also via female-to-female jumper wires. This connection will power the Raspberry Pi 4B which will be the “brain” of the GPMS.

As the Pi is powered up and also taking in image data, the resulting output will be the input to the projector via HDMI. The Raspberry Pi 4B will be in close proximity to the custom PCB with plans to design a hardware enclosure to house the two boards for a touching simplicity and portability aspect. [49]

After initial PCB manufacturing and extensive testing the authors decided to perform minor adjustments to the first schematic which can be seen in Figure 6.2.2. A comprehensive explanation regarding the reasons behind a new schematic are expanded upon in Section 9.1.3.

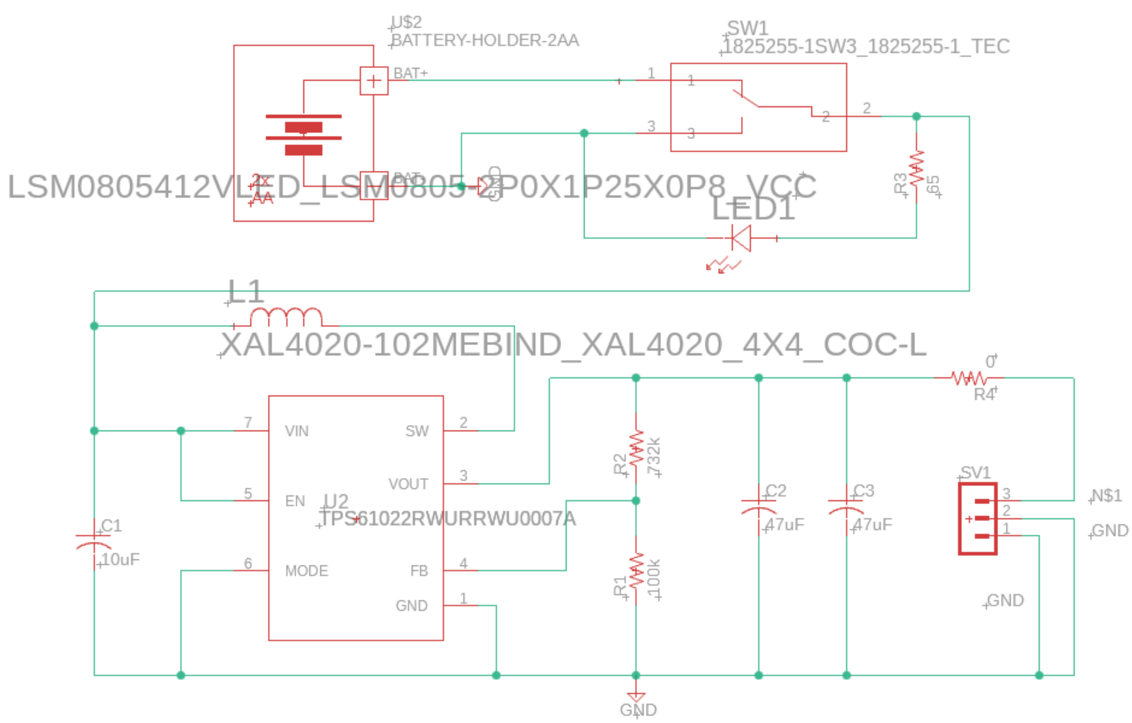


Figure 6.2.3 *Custom PCB (Second Revision)*

As shown above, there is now only one battery holder. This was done in order to prevent battery drainage when two battery holders are connected in parallel. After collaboration with Dr. Weeks of the University of Central Florida, it was determined that placing batteries in parallel on the first iteration of the custom PCB presented several significant drawbacks. Firstly, even minor voltage differences between batteries can lead to imbalanced charging and discharging, reducing the overall efficiency and potentially shortening battery lifespan. This was imperative given that the entire system essentially relies on the Raspberry Pi 4B having a consistent PSU.

Additionally, variations in battery capacity can result in uneven distribution of load, causing premature failure of weaker cells. As a result, managing multiple parallel batteries was deemed a proven roadblock to efficiency and an increase to design complexity. Therefore, avoiding parallel battery configurations on the second schematic of the custom PCB was crucial to ensure optimal efficiency and longevity of the battery system.

In this second iteration, the negative terminals of each battery holder and *now* the LED are directly connected to ground. Similar to the first revision of the custom PCB which can be seen in Figure 6.1.1, when the switch is turned *off*, no current will flow through the switch. When the switch is turned *on*, current will first flow through to a resistor to limit itself before reaching the red LED. When the current reaches the red LED, it will turn on to indicate it is on. The current will also pass through a DC/DC converter to up the voltage to 5V at 3A, which is what the Raspberry Pi 4B requires. [47] There is also a 0Ω resistor at the output of the regulator to make further troubleshooting the power supply much easier. The output of the DC/DC converter still goes to a 3 pin header with the top most pin being the pin that connects to the 5V rail on the Raspberry Pi 4B. The middle and bottom most pins remain connected to ground. [13]

6.3 PSU Time Specifications

Given the importance of the role of the custom PCB in GPMS, it is important to accurately calculate the total time specifications so that the authors know how long GPMS can last before having to replace the batteries.

To ascertain the total energy capacity of the batteries powering the Raspberry Pi 4B, the authors first converted the board's power consumption to Watts which will be shortly used in Figure 6.3.4.

$$Power\ (W)\ =\ Voltage\ (V)\ *\ Current\ (A)$$

Figure 6.3.1 *Equation for Power*

$$Power\ (W)\ =\ 5V\ *\ 3A\ =\ 15W$$

Given the multiplication of 5V and 3A, the Pi's power consumption yielded a total of 15W. The significance of using only one single battery holder lies in maintaining a consistent voltage of 3.3V. Since the batteries are in series, when the voltages of each battery (1.65V) are added together, they yield a total voltage of 3.3V. Furthermore, the capacity of 2.7Ah per battery remains the same since both batteries are in series. The energy output of one set of batteries, with two batteries in series, can be calculated using the following equation:

$$Energy(Wh) \; = \; Voltage\;(V) \; * \; Capacity\;(Ah)$$

Figure 6.3.2 *Equation for Energy*

$$Energy(Wh) \; = \; 3.3V \; * \; 2.7AH \; = \; 8.91\,Wh$$

Given the voltage remains at 3.3V, the energy output per set of batteries is 8.91Wh. Furthermore, the revised total energy output serves as the basis for calculating the operational duration of the Raspberry Pi 4B, given its power consumption rate of 15W. Utilizing the formula:

$$Hours\ of\ Operation \; = \; \frac{Total\ Energy\ (Wh)}{Power\ Consumption\ of\ Raspberry\ Pi\ 4(W)}$$

Figure 6.3.3 *Equation for Hours of Operation*

$$Hours\ of\ Operation \; = \; \frac{8.91\,Wh}{15\,W}$$
$$Hours\ of\ Operation \approx \; 0.6\ hours$$

The authors therefore compute the operational duration to be approximately 36 minutes which is adequate for the early stages of development of GPMS. Provided that the Raspberry Pi 4B requires an input voltage of 5V at 3A, it is important to note that the 3A specification is with regards to all peripherals in use consuming near maximal power. Therefore, the duration of operation of the Raspberry Pi 4B is likely to last closer to the 1 hour target specification mentioned in the House of Quality shown in Figure 2.4.6.

6.4 Overall Schematic

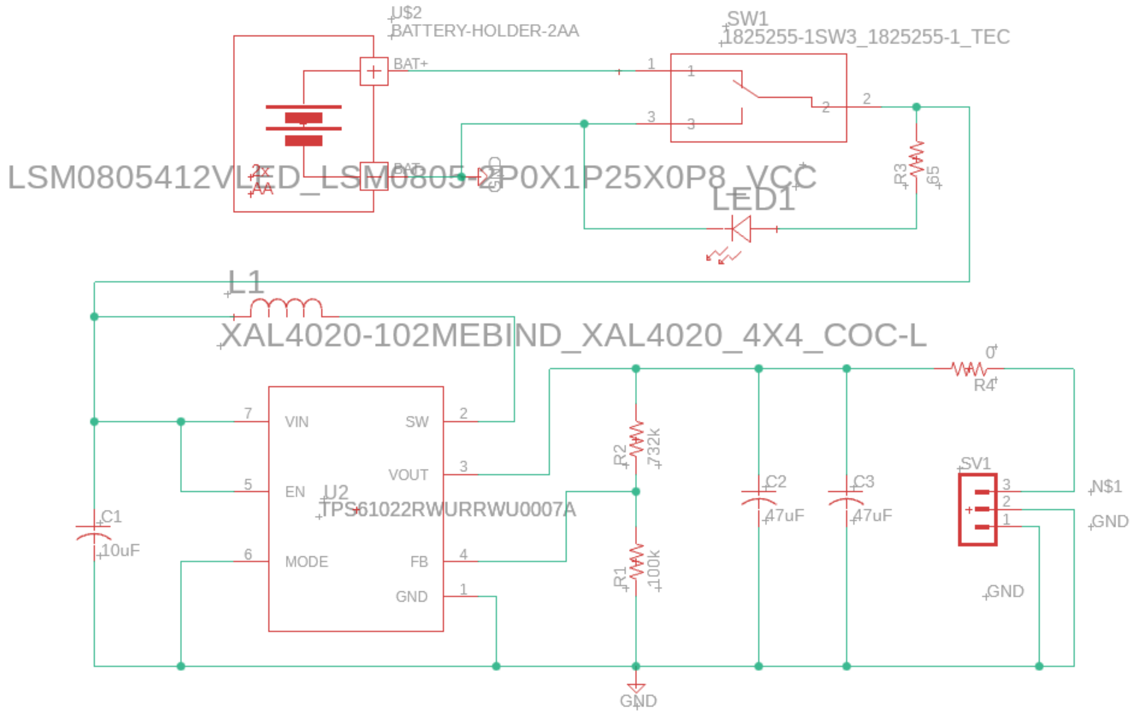


Figure 6.4 *Schematic for System*

As demonstrated in the above schematic, the 5V+OUTPUT net via pin 3 on the header, is to be connected to the V_{in} pin of the Raspberry Pi 4B, and both of the GND pins to the corresponding ground pins on the Pi as well which can both can be located Figure E. This Raspberry Pi is responsible for controlling the substantive portion of the design, connecting the device via micro-HDMI to both the projector as well as the touch screen. Additionally, the display will receive power and transmit touch data via a single USB connection with the Raspberry Pi 4B.

Chapter 7 Software Design

7.1 Software Design Overview

Due to the nature of GPMS and its number of off-the-shelf hardware features, cohesive software design is paramount to providing the users with a tailored experience that aligns with the authors' vision. As demonstrated in Section 2.5.3, *Software Diagram*, there are 2 major platforms that will need to be designed and developed for: the GPMS device and the generative server stack.

In this chapter, the authors will discuss each of these design platforms in detail, highlighting the flow of data both internally and between platforms, as well as the interaction between the user and these unique systems.

7.2 Compute on Device

Within the framework of this system, the Raspberry Pi 4B serves as the central processing unit, ensuring the seamless integration of hardware inputs and user directives. Its role begins at the very onset of the device's activation, handling the initialization that calibrates the camera to correct for lens distortion and aligning the projection parameters with the user’s environmental context. This ensures that the images produced are tailored to the physical constraints of the projection surface. The Pi's computation is also leveraged in providing a user-friendly interface through a 10" LCD capacitive touch display, allowing users to input their creative prompts and adjust projection settings in real time.

As the device transitions to its main operational phase, the Raspberry Pi maintains its pivotal role by managing the “main loop”—a continuous cycle of user interaction, data processing, and projection adjustments. It handles the intricacies of network connectivity, packaging user inputs and sensor data, and dispatching them to an external server for processing.

The primary tools/languages to be used in developing this platform are OpenCV for its extensive open-source library of computer vision processes, Qt for its native front-end framework, and C++/Python to integrate these tools into a cohesive, full-stack device.

7.2.1 Camera Calibration

The initialization process is a critical sequence that ensures the Raspberry Pi, in conjunction with the OV5640 camera, is correctly calibrated for projecting images. As the device powers up, it immediately engages in a pre-calibration routine to counteract the typical distortions produced by the camera's lens. This step is automated, using the camera's specifications to apply a real-time correction that flattens and aligns the image edges with the projector's target surface.

In the camera calibration process using OpenCV, the focus is on correcting lens distortion to enhance the accuracy of image projection. This involves compensating for both radial and tangential distortions. Radial distortion, which creates a "barrel" or "fish-eye" effect, is addressed using a formula that recalculates pixel positions based on the distortion coefficients. Tangential distortion is corrected through a similar set of formulas, adjusting for the misalignment of the lens with the imaging plane. Both types of distortions are represented in OpenCV by a matrix, which contains five coefficients, $[k_1, k_2, k_3, p_1, p_2]$, demonstrated by the equations below.

$x_{distorted} = x(1 + k_1r^2 + k_2r^4 + k_3r^6)$ $y_{distorted} = x(1 + k_1r^2 + k_2r^4 + k_3r^6)$	$x_{distorted} = x + [2p_1xy + p_2(r^2 + 2x^2)]$ $y_{distorted} = y + [p_1(r^2 + 2y^2) + 2p_2xy]$
Figure 7.2.1.1 *Equations for Radial Factors*	**Figure 7.2.1.2** *Equations for Tangential Distortion*

Table 7.2.1.1 *Distortion Equations* [50]

In addition to distortion correction, the calibration process includes determining the camera matrix that translates between camera coordinates and real-world units. The camera matrix accounts for the camera's focal lengths and the optical centers expressed in pixel coordinates. For reliable calibration, OpenCV can utilize various types of patterns, such as chessboard or circle grids, captured in multiple snapshots from different angles. These snapshots allow OpenCV to formulate a set of equations that are solved to determine the camera matrix and distortion coefficients. Once the parameters are obtained, functions like cv::undistort can be used to correct the captured images, significantly improving the quality of the projected image even for low-cost and low-quality cameras. [50] Lastly, it's important to note that these 5 coefficients can be pre-computed using the exact camera from GPMS and saved to the device for later use, allowing this tuning to take place instantly without any user interaction.

The importance of this process can be seen below in the previously discussed work of Addison Sandvik, who provides images showcasing the importance of lens correction. Note the roundness of the edges towards the horizontal and vertical bounds of the image.

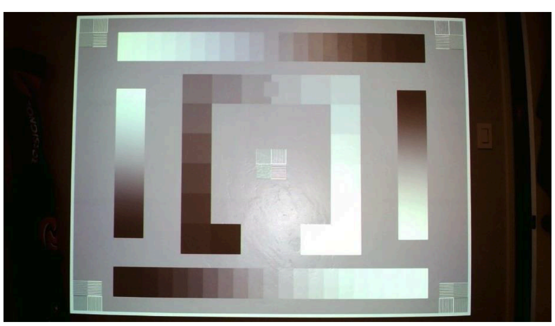
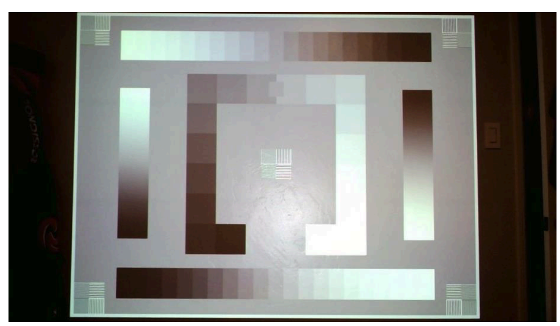
	

Figure 7.2.1.3 *Without Lens Correction* [4]	**Figure 7.2.1.4** *With Lens Correction* [4]

Table 7.2.1.2 *Comparison Using Lens Distortion* [4]

Following this automated correction, the system transitions to a user-interactive mode. It projects a calibration image onto the surface, which can be used to select the 4 outer corners of the display (i.e. the corner points that define the boundaries of the projection area). The 10" LCD capacitive screen is used to provide a tactile interface for the user. They can directly interact with the calibration image, dragging and adjusting corners to achieve an exact overlay of the projected image with the physical boundaries. Next, using a transform, the image capture is cropped/shifted using these points to account for unevenness in the edges. This careful alignment is what ensures that subsequent images are projected with precision, fitting neatly within the desired area. From here, edge detection is completed to show the aligned vertices.

7.2.2 Threshold Calibration

Another essential aspect of initialization is defining the threshold values for the edge detection process. These cannot be automatically set without complex computer vision algorithms due to any number of environmental variables the GPMS might be introduced to (i.e., lighting, diverse colors, patterns, etc.); therefore, it is preferable for the user to set these values manually using a simple slider to decide what looks best.

These thresholds act as filters determining what constitutes an edge, thereby defining the critical contours of the object/shape. Lower threshold values may result in the detection of faint or less pronounced edges, potentially introducing noise and unwanted details. Conversely, excessively high threshold values can lead to the omission of significant edges, rendering the object unrecognizable. The fine line between these extremes is where the optimal edge definition is achieved, ensuring that the object is accurately captured without superfluous data.

Below is an example using an early prototype of the project to show the differences between various threshold values and the importance of the user fine-tuning these values manually. The image used to highlight these differences is *“How to Draw Pusheen with Donut”* [51].

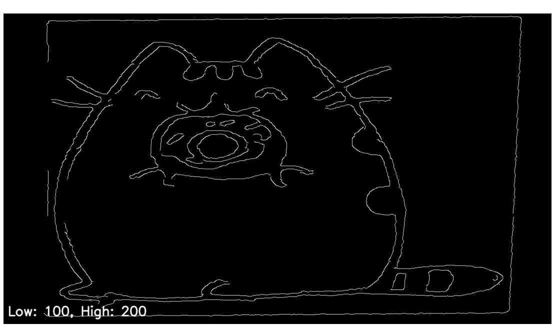


Figure 7.2.2.1
Thresholding - Low: 100, High: 200

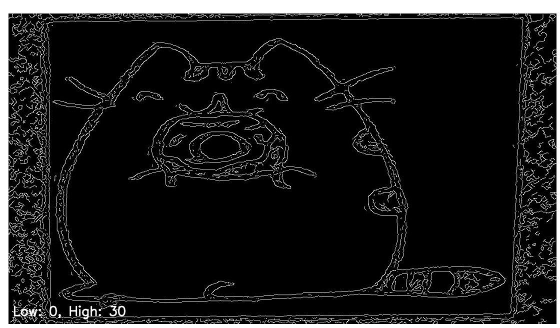


Figure 7.2.2.2
Thresholding - Low: 0, High: 30

 **Figure 7.2.2.3** *Thresholding - Low: 0, High: 10*	 **Figure 7.2.2.4** *Thresholding - Low: 0, High: 0*

Table 7.2.2 *Thresholding Comparison*

In the first image, with a low threshold of 100 and a high of 200, moderate edge detection is observed, with key boundaries of the subject discerned, yet some finer details are potentially lost. This setting might be ideal for scenarios where the primary goal is to detect prominent features while avoiding excessive noise that could complicate the analysis.

The subsequent images, beginning in the top right, depict the increasing impact of lowering both the high and low threshold values, with the last one featuring the most extreme case where both thresholds are set to 0. The progressive reduction in thresholds results in a more comprehensive capture of edges, including more subtle variations in the image. However, this sensitivity also includes undesirable noise, highlighting the critical nature of striking a balance between detail and clarity.

7.2.3 Main Loop

Once initialization concludes, the main loop begins. The 10" LCD is the interactive gateway for users to input their creative directives. The system is designed for straightforward engagement: users can type/select prompts and choose from various styles for how the image should look.

The most important aspect of this step is that the UI is clean and responsive. The design of the UI/UX will be discussed further in section 7.4.

After entering a prompt and selecting a style, the user will be presented with a choice of 4 images matching their creative aspirations. They will be able to choose any one of these, save them for later display, or regenerate a new set of images. From here, when they have found an image they think is suitable, they will be able to then display it on the given surface by selecting it. Finally, when they want to move the project or adjust it for any reason, they will also be able to re-enter the initialization process at any time to re-calibrate the frame.

7.2.4 Network Connectivity

Next, the camera data, alongside the user's input and selected threshold values, are bundled and transmitted to a non-local server for processing using an SSH connection powered by TCP to generate the final image before projection on the surface.

This process will be handled by a Python script that establishes the SSH connection, handles the requisite SCP commands, and is able to trigger the server's Python script remotely.

This connection utilizes a non-traditional method to connect the components responsible for the data transfer/image request through the available IP address on the university network. The data packet is made up of 2 critical aspects, which are added to a folder and zipped before transferring via SCP to the server (a protocol established through SSH). These files will be made up of the following information:

info.json
- Threshold Lo: uint8_t
- Threshold Hi: uint8_t
- Prompt: string
- Style: string
- Image Size (4 bytes): uint32_t

Image.png
- The image of the scene that can be used to map ControlNet

From here, the script will again utilize SCP to access the newly created files, downloading them onto the Raspberry Pi for selection and display.

It is important to note the context in which this design has been engineered. The machine being used exists in a secure UCF Lab on campus, as AMD has provided the authors access to it. The computer has only been made available remotely through SSH, and there is an unlikely chance of UCF allowing a port to be made available to create a legitimate REST API. If this project were to become its own legitimate venture outside of academia, this network architecture would be changed to a more normalized networking standard, likely using some sort of formal API process on an available port. These constraints are further discussed in Section 4.2.4, *Constraints in Security/Network Integration*.

7.3 User Interface

The user interface of the Graphical Projection Mapping System (GPMS) is designed to provide a seamless and intuitive experience for users, enabling them to generate and display images on physical structures with ease. The authors have conducted extensive research on application design principles to create an interface that is both user-friendly and visually appealing, focusing on simplicity, clarity, and consistency.

The main objectives of the GPMS user interface are to guide users through the process of capturing an image, adjusting edge detection settings, inputting a text prompt, selecting a generated image, and aligning the image with the physical structure. By minimizing visual clutter and providing interactive controls, the interface aims to enhance the user experience and support the core functionalities of the application.

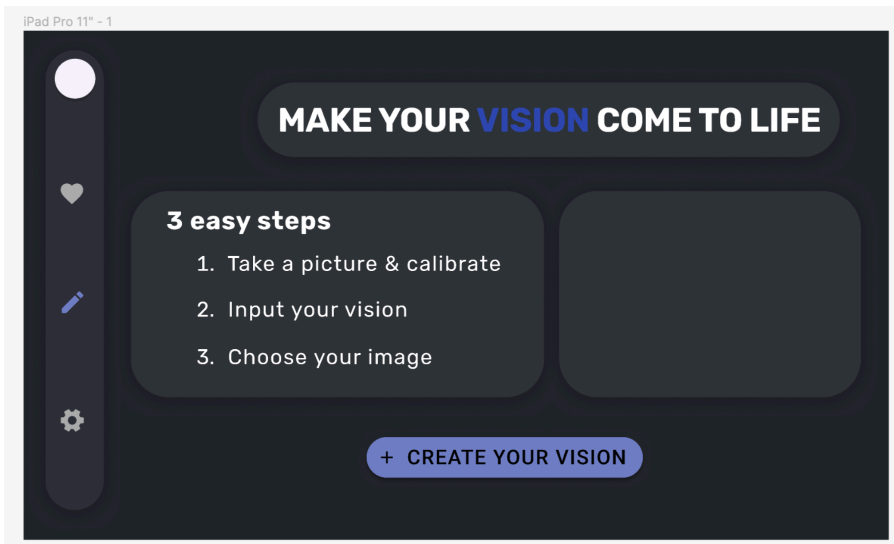


Figure 7.3.1 *Home Page*

The workflow of the GPMS application begins on the home page (Figure 7.3.1), where users can initiate the image capture process. Upon clicking the "Submit" button, users are directed to the picture page (Figure 7.3.2), where they can approve the captured image or choose to retake the picture if needed. This flexibility allows users to ensure that the captured image accurately represents the desired structure.

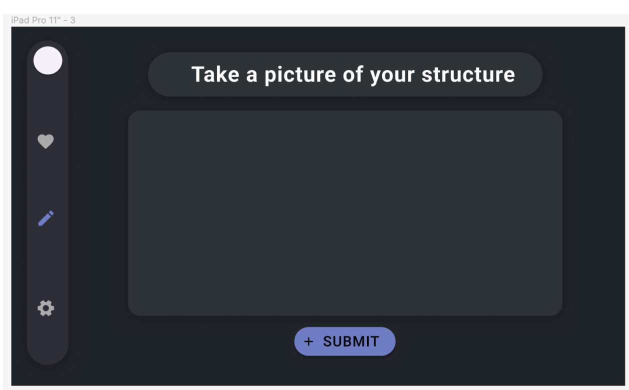


Figure 7.3.2 *Picture Page*

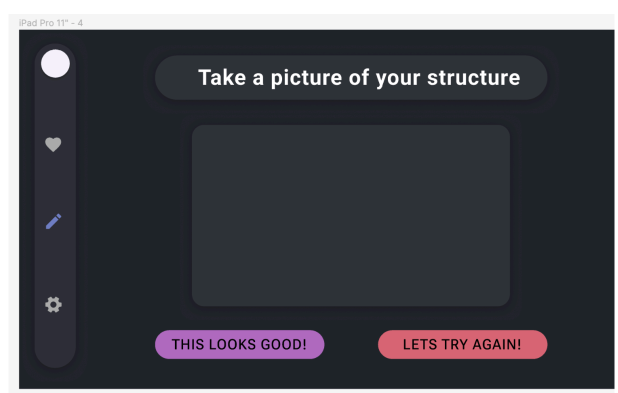


Figure 7.3.3 *Picture Approval Page*

Once the image is approved, the application displays an edge-detected version of the image on the sensitivity threshold page (Figure 7.3.4). Here, users can interact with a slider to adjust the amount of edges included in the image, providing control over the level of detail. This feature enables users to fine-tune the edge detection process to suit their specific needs and preferences.

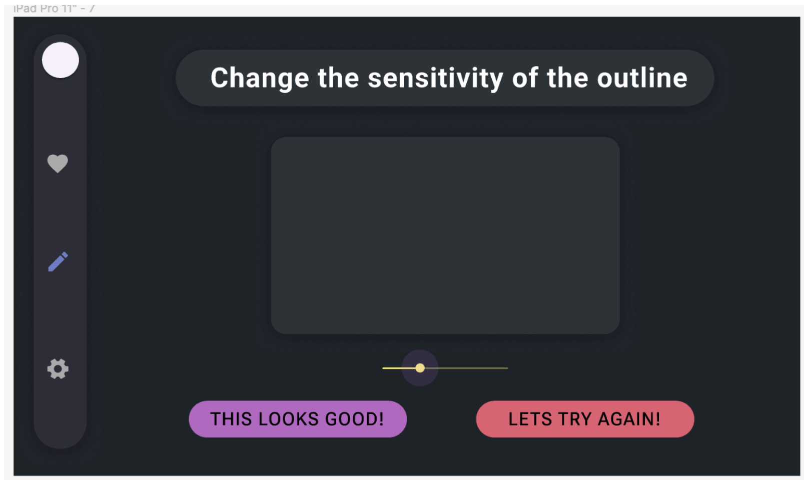


Figure 7.3.4 *Sensitivity Threshold Page*

After finalizing the edge detection settings, users proceed to the input prompt page (Figure 7.3.5), where they can enter a text prompt to guide the image generation process. The application then utilizes both the user-provided prompt and the captured structure to generate a set of four images, which are presented on the generated image selection page (Figure 7.3.1). Users can browse through these images, save their favorites, and select the image they wish to display on the physical structure.

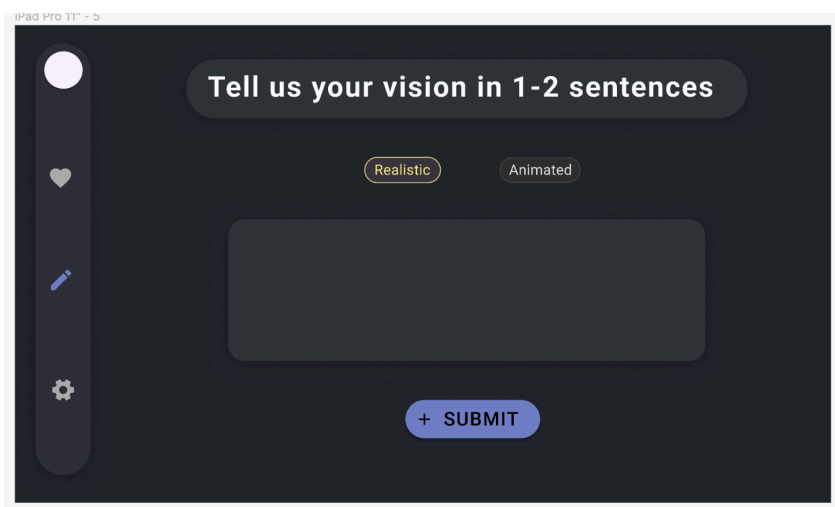


Figure 7.3.5 *Input Prompt Page*

To ensure accurate projection of the selected image onto the structure, the application provides an intuitive cropping-like function. Users can drag the corners of the image to their corresponding locations on the physical structure, allowing for precise positioning and scaling. This feature is essential for achieving a seamless and visually coherent projection, aligning the edges of the image with the real-world structure.

Throughout the development process, the authors prioritized design principles such as simplicity, clarity, and consistency. The user interface features a dark theme, which not only reduces visual clutter but also enhances the viewing experience in low-light environments. The minimalist design approach ensures that users can focus on the core functionalities of the application without unnecessary distractions.

In addition to its core features, the GPMS user interface incorporates personalization options, such as the ability to save and "favorite" generated images. This allows users to curate their own collection of artworks and easily access their preferred images for future use.

The GPMS user interface offers a compelling and user-friendly experience, guiding users through the process of generating and displaying images on physical structures. By combining intuitive controls, interactive features, and a visually appealing design, the interface supports the core functionalities of the application while prioritizing simplicity and clarity. As the project evolves, the authors will continue to gather user feedback and explore potential improvements to further enhance the user experience.

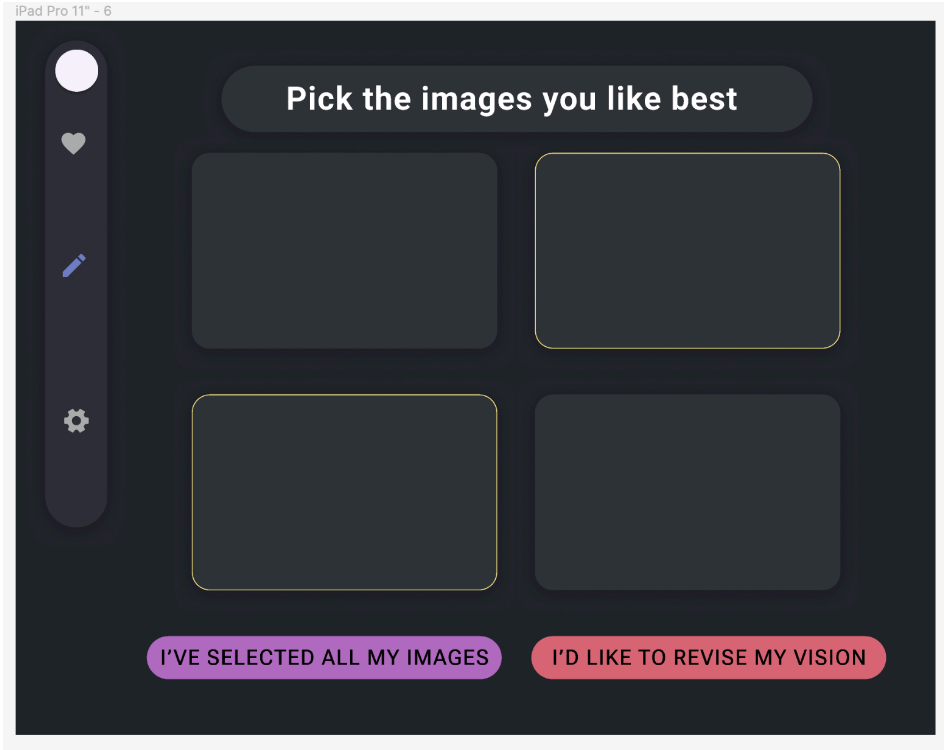


Figure 7.3.6 *Generated Image Selection Page*

The user interface of GPMS was meticulously prototyped using Figma, a powerful wireframing and design collaboration tool. By leveraging Figma's capabilities, the team ensured that all members had a clear understanding of the application's flow, visual aesthetics, and user experience. This collaborative approach allowed for seamless communication and alignment among the team, fostering a unified vision for the project.

During the initial stages of the design process, the team considered implementing a vibrant and cheerful color scheme to create an engaging and lively user interface. However, a crucial insight from one of the team members prompted a significant shift in the design direction. It was astutely pointed out that GPMS would likely be used primarily in dimly lit or dark settings. In these environments, a bright and colorful interface could potentially interfere with the immersive experience and distract users from the core functionality of the application.

Taking this valuable feedback into account, the team made a strategic decision to adopt a darker themed user interface. By employing a color palette dominated by darker shades and muted tones, the interface would seamlessly blend into the ambient lighting conditions, minimizing visual distractions and allowing users to focus on the primary tasks at hand. This design choice not only enhances the overall user experience but also demonstrates the team's adaptability and willingness to incorporate insights that align with the specific context and requirements of the application.

The decision to pivot towards a darker themed UI exemplifies the team's commitment to creating an interface that is not only visually appealing but also functionally optimized for the intended use case. By carefully considering the environmental factors and user needs, the team has crafted a user interface that strikes a harmonious balance between aesthetics and practicality, ultimately enhancing the usability and effectiveness of the GPMS application.

7.4 Generative Design Stack

This section focuses on the image generation process that creates images based on the structure's edges and the user's prompt. The process takes input from the user interface, which is passed through a script (as mentioned in Section 7.2.4) and then to a program responsible for generating the images. The program's process and functionalities will be explained in detail in this section.

The image generation process involves several key components, including prompt expansion, edge detection pre-processing, ControlNets, and the Stable Diffusion model [52]. These components work together to generate images that accurately reflect the user's prompt and incorporate the structure's edge information.

First, the user's prompt undergoes prompt expansion, a technique that enhances the accuracy and detail of the generated images. Next, the structure's edge information is obtained through the image taken by the user through the user interface. The image is passed through a specific edge detection preprocessor to extract the relevant edge features.

The processed edge image is then fed into a ControlNet, a neural network that allows for more user control over the image generation process. The ControlNet outputs a set of conditioning vectors or maps that encode the desired modifications to the Stable Diffusion model. The Stable Diffusion model is a deep learning model trained to generate images based on input prompts and conditioning information. It uses the conditioning vectors or maps from the ControlNet and the expanded prompt to guide the image generation process, iteratively refining the output image. The following subsections will delve into each component of the image generation process in more detail.

7.4.1 Prompt Expansion

Prompt expansion is a technique that plays a crucial role in generating more detailed and vivid images when working with Text-to-Image models. It addresses the common challenge of users providing insufficient information in their input prompts, which can result in less descriptive and less accurate generated images. By employing prompt expansion, the system can enhance the user's input prompt by adding more relevant details and context.

During the research phase, it became apparent that while there are numerous models available for Text-to-Image generation, finding a suitable model specifically designed for prompt expansion proved to be more challenging. The authors encountered difficulties in identifying a free and readily available model that was tailored for Text-to-Image generation. It is crucial to select a prompt expansion model that is compatible with Text-to-Image tasks, as using a model intended for other purposes may potentially degrade the quality of the generated images.

After extensive research, the authors discovered that one viable approach is to develop a custom prompt expansion model. While the Stable Diffusion model itself incorporates a built-in prompt expansion mechanism, user feedback suggests that employing an additional prompt expansion step can further improve the quality and coherence of the generated images [52]. However, before investing resources into creating a custom prompt expansion model, the authors will need to assess the performance of the Generative Projection Mapping System (GPMS) without prompt expansion. This evaluation will help determine whether prompt expansion is indeed necessary for achieving the desired results.

If the assessment indicates that prompt expansion would significantly enhance the system's output, or if time permits, the authors will further investigate the process of developing their own prompt expansion model. This may involve exploring techniques such as fine-tuning pre-trained language models, leveraging domain-specific datasets, or employing few-shot learning approaches. Alternatively, the authors will continue their search for a suitable prompt expansion model that is specifically designed for Text-to-Image generation. This may involve collaborating with the research community, exploring open-source repositories, or considering commercial solutions that align with the project's requirements and constraints.

Incorporating an effective prompt expansion module into the GPMS pipeline has the potential to greatly improve the quality, relevance, and creativity of the generated images. By providing more comprehensive and detailed prompts, the system can better capture the user's intent and generate images that are more visually appealing and semantically meaningful.

7.4.2 Edge Detection Pre-Processing

As discussed in the introduction of this section, pre-processing must occur on the image prior to being inputted into the ControlNet. This pre-processing step involves applying specific edge detection techniques to the image and returning a new image that highlights the relevant edge features. Edge detection is a crucial component in the Stable Diffusion with ControlNets pipeline, as it provides the necessary conditioning information for the ControlNet to guide the image generation process.

There are several types of edge detection techniques available for Stable Diffusion with ControlNets. In this project, we have access to eight different edge detection models: canny, depth, HED boundary, scribble, semantic segmentation, normal, Openpose, and M-LSD. Each of these models has its own unique characteristics and outputs, as shown in the table below. The example images are sourced from the GitHub repository that contains the respective models.

Edge Detector	Input Image	Output Image
Canny	 **Figure 7.4.1.1** *Pre-canny*	 **Figure 7.4.1.2** *Post-canny*
Normal	 **Figure 7.4.1.3** *Pre-Normal*	 **Figure 7.4.1.4** *Post-Normal*

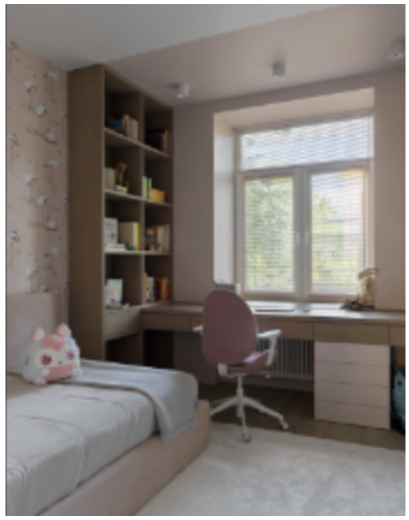
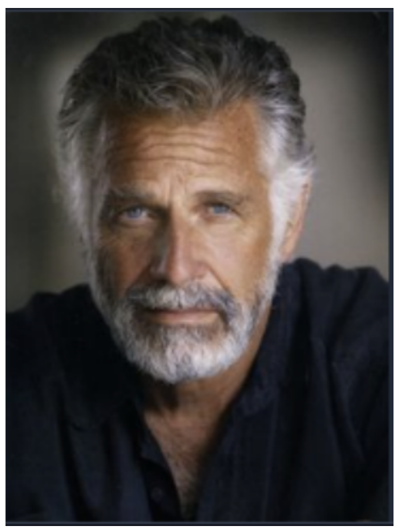
M-LSD	 **Figure 7.4.1.5** *Pre-M-LSD*	 **Figure 7.4.1.6** *Post-M-LSD*
HED	 **Figure 7.4.1.7** *Pre-HED*	 **Figure 7.4.1.8** *Post-HED*

Table 7.4.2 *Edge Detection Models* [7] [53]

The choice of edge detection model depends on the specific requirements and characteristics of the project. In the context of this project, where the projector is expected to be positioned approximately 10 feet away from the structure, certain models may be more suitable than others.

Based on the distance and the nature of the structure, models like semantic segmentation and depth may not be the most effective. These models tend to perform better when capturing details at further distances. On the other hand, edge detection models such as canny, M-LSD, and HED appear to be the most promising candidates for this project.

The type of edge detection model used in this step can have a significant effect on the output of the ControlNet and therefore the final generated image. Different edge detection models capture different types of edge information and provide varying levels of detail, which determines how ControlNet regulates the Stable Diffusion through the image generation process. Therefore edge characteristics and the level of detail are important characteristics to make note of when choosing a model.

Canny edge detection is a widely used technique that identifies edges by looking for the local maxima of the gradient of the image. It applies noise reduction, non-maximum suppression, and hysteresis thresholding to produce clean and accurate edge maps. Canny edge detection is known for its ability to detect fine details and produce sharp edges. Canny edge detection models focus on capturing fine-grained edges and produce thin, precise edge maps. This can lead to generated images with sharp and well-defined edges, which can enable the ControlNet to guide the Stable Diffusion model towards generating images with more precise and detailed features. [54]

M-LSD (Multiscale Line Segment Detector) is a line segment detection algorithm that identifies meaningful line segments in an image. It is particularly effective in capturing long and straight edges, making it suitable for detecting the overall structure and boundaries of objects. This can result in generated images with more pronounced and coherent structural elements. Because this model focuses on higher-level edge information it may result in generated images with more abstract or simplified representations of the structure. [55]

HED (Holistically-Nested Edge Detection) is a deep learning-based edge detection model that leverages a fully convolutional neural network architecture. It is trained to predict edges at multiple scales and combines the predictions to produce a final edge map. HED is known for its ability to capture both fine-grained and high-level edge information. This can lead to generated images with a balance of local and global edge coherence, which means the image generated is more precise and with detailed features, similar to the canny edge detection model. [56]

The selection of the appropriate edge detection model will be based on empirical evaluation and testing with multiple structures and lighting conditions of the project. We will conduct experiments and assess the performance of each model, considering factors such as the accuracy of edge detection, the level of detail captured, and the overall quality of the generated images. Through these evaluations, we will determine which edge detection model or models are best suited for our specific use case. It is possible that a single model may emerge as the most effective choice, or we may find that different models excel in different scenarios.

Furthermore, we will explore the possibility of providing users with the option to choose their preferred edge detection model. This would allow for greater flexibility and customization, enabling users to select the model that aligns with their specific requirements or artistic preferences. By offering a choice of edge detection models, we can cater to a wider range of user needs and enhance the overall user experience.

Ultimately, the decision on which edge detection models to incorporate into the final system will be made based on a combination of performance evaluations and the specific goals of the project. The pre-processed edge image, obtained from the selected model, will serve as the conditioning input for the ControlNet, enabling it to guide the Stable Diffusion model towards generating images that align with the desired structure and edge features.

7.4.3 Stable Diffusion with ControlNet

This part of the process runs concurrently with the edge detection pre-processing. As explained in detail in Section 2.2.4 Stable Diffusion with ControlNet, ControlNet is essentially a copy of the Stable Diffusion model that functions as a HyperNetwork. It injects weights from ControlNet into the Stable Diffusion model, allowing information from the pre-processed image to be incorporated into the final output.

In this project, we will utilize the Stable Diffusion with ControlNet Pipeline from the Diffusers library provided by Hugging Face. The pipeline employs the same ControlNet models created by the original ControlNet developers. These models include both the Stable Diffusion model and ControlNet, along with the specific edge detection model used. It's important to note that ControlNet is also tailored to the particular edge detection model employed.

By using the pipeline, we can easily import the appropriate model and input all the necessary user information, which is transmitted through the zip file as described in Section 7.2.4. The pipeline handles the integration of the pre-processed edge image with the ControlNet and Stable Diffusion models, generating the final output image. The generated image is then sent back to the Raspberry Pi using the same communication mechanism outlined in Section 7.2.4.

The Stable Diffusion with ControlNet Pipeline simplifies the implementation process by providing a high-level interface for utilizing ControlNet models. It abstracts away the complexities of integrating ControlNet with Stable Diffusion, allowing us to focus on providing the necessary inputs and handling the generated output. This streamlined approach enhances the efficiency and maintainability of our image generation pipeline.

7.4.4 The Generate Program

The entire image generation process will be encapsulated within a single program, which will be executed by the script described in Section 7.2.4. This program will follow the sequence of steps outlined in Section 7.3, with all the required variables being passed through the script. The program will begin by applying a prompt expansion model to ensure that the user input prompt is as accurate and detailed as possible. This step aims to enhance the quality and relevance of the generated images by providing a more comprehensive and specific prompt.

Next, the program will determine which edge detection model to use based on the results of our testing and evaluation, as explained in Section 7.3.2. Depending on our findings, the edge detection model may be pre-selected by the authors or chosen by the users through the user interface. Once the appropriate edge detection model is determined, the program will proceed to pre-process the input image using the selected model. This step involves applying the specific edge detection technique to extract the relevant edge features from the image. With the pre-processed edge image available, the program will then invoke the corresponding Stable Diffusion with ControlNet model, which is specifically associated with the chosen edge detection model. The ControlNet model will guide the image generation process by conditioning the Stable Diffusion model based on the edge information provided.

The Stable Diffusion with ControlNet model will generate the final output image’s based on the user's prompt and the pre-processed edge image. These generated images will be returned to the Raspberry Pi and displayed on the user interface for the user to view and interact with. By encapsulating the entire image generation process within a single program, we ensure a cohesive and efficient workflow. The program handles the prompt expansion, edge detection model selection, pre-processing, and integration with the Stable Diffusion and ControlNet models. This modular approach allows for easy maintenance, debugging, and future enhancements to the image generation pipeline.

Overall, the Generate Program serves as the central component that orchestrates the various steps involved in generating images based on user prompts and edge detection techniques. It leverages the power of Stable Diffusion with ControlNet to produce visually compelling and structurally coherent images that align with the user's intentions.

Chapter 8 System Fabrication/Prototype Construction

8.1 Custom PCB Design

The first iteration of the custom PCB is designed via AutoDesk Fusion360. The following custom PCB includes both grounding and polygon outlining to distinguish a secure connection between the top and bottom layers.

The custom PCB without the top and bottom layer pour is listed below in Figure 8.1.1, with via's visible and strategically placed, demonstrating that the top and bottom board layers are connected. The air wires were also routed so that no wires overlapped, hence the red-colored wires. The only instance of overlap was solved when the overlapping wire was routed on the second layer, hence the blue color wire.

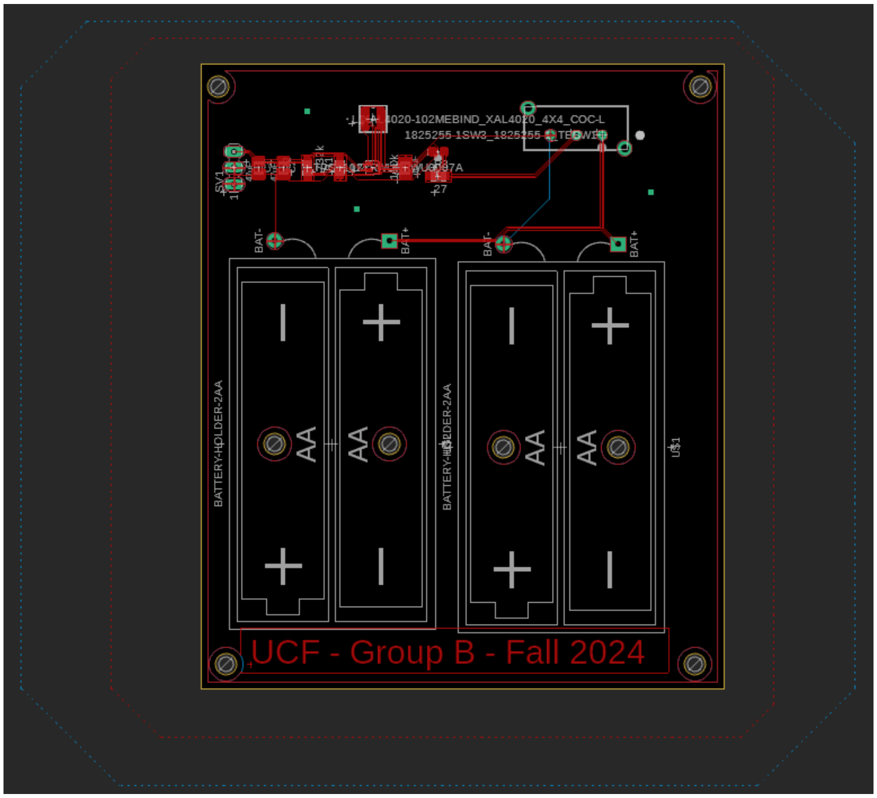


Figure 8.1.1 *Custom PCB Without Pour (First Revision)*

The first iteration of the custom PCB with both the top and bottom layer pour is shown below in Figure 8.1.2.

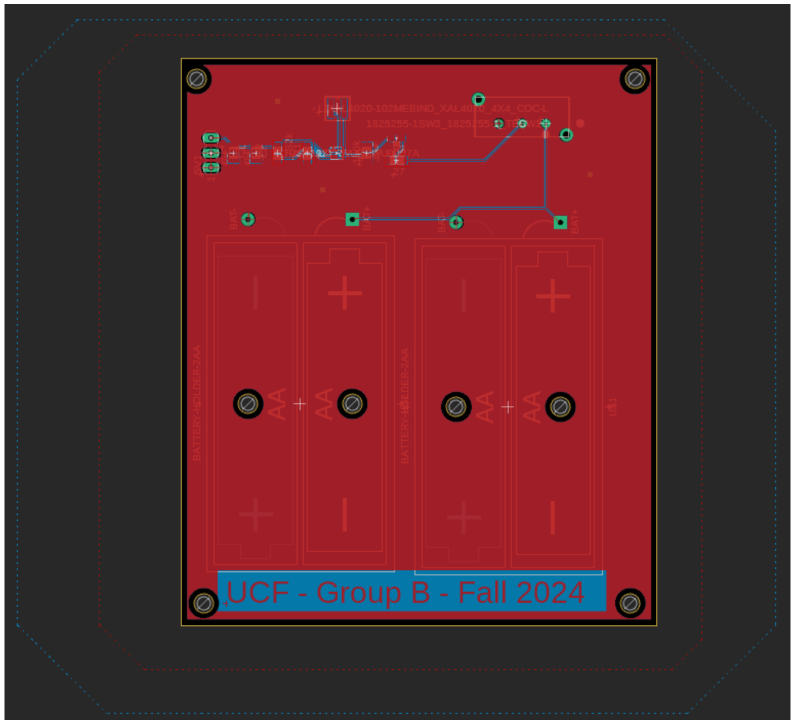


Figure 8.1.2 *Custom PCB With Pour (First Revision)*

The first iteration of the custom PCB in its 3D version is listed below in Figure 8.1.3. The 3D model of the custom PCB purposely and precisely resembles the schematic seen in Figure 6.2.1.

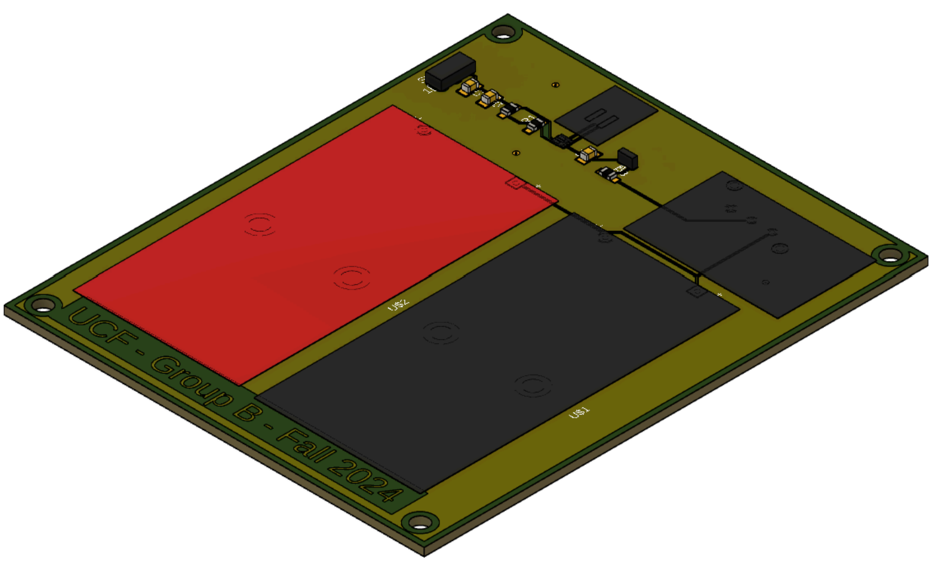


Figure 8.1.3 *3D Custom PCB (First Revision)*

Overall, the custom PCB design ensures clarity and ease of comprehension for reviewers. At first, it was believed that efficient power management is achieved by utilizing two Alkaline battery holders and a DC/DC converter, delivering a stable 5V supply to the Raspberry Pi 4B while powering additional components. User experience, a vital theme to GPMS, is enhanced with user-friendly controls, including an on/off switch with an indicator LED that provides clear visual feedback and enhances awareness.

LED protection is assured through calculated resistor values, safeguarding against excessive current and preventing burnout. The design easily accommodates diverse components by offering versatile power supplies catering specifically to varied voltage requirements. A convenient interface utilizing female-to-female jumper wires streamlines integration and maintenance, enabling seamless communication and power distribution. Finally, compact integration by closely approximating the distance between the Raspberry Pi 4B directly next to the custom PCB optimizes space and assembly, refining the overall system form factor. In addition, a hardware enclosure will be manufactured eventually throughout Senior Design 2.

This approach combines meticulous planning, efficient power management, user-friendly controls, and careful component selection to create a comprehensive and efficient custom PCB design for powering and interfacing with a Raspberry Pi 4B, suitable for GPMS. Additionally, the custom PCB will be ordered from JLCPCB, given the standout features that JLCPCB offers [29] [57].

One of the standout features of JLCPCB is its cost-effectiveness. Known for competitive pricing, particularly for small orders, JLCPCB often runs promotions and discounts. Notably, they offer a popular five boards for 2 dollars deal for simple PCBs, which aligns perfectly with the authors' needs and what they took advantage of. In addition to cost-effectiveness, JLCPCB boasts fast turnaround times. They can produce simple PCBs in as little as 24 hours. This rapid turnaround will be crucial for the authors' prototyping needs, enabling them to move swiftly with their project advancements.

Furthermore, JLCPCB provides an easy-to-use online platform. Uploading Gerber files and placing orders is straightforward with a user-friendly web interface. This simplicity and convenience are especially beneficial for the authors, who are relatively new to PCB manufacturing processes

The top production file image is listed below in Figure 8.1.6. It represents the first manufactured physical iteration of the custom PCB that will arrive at the authors. In addition to the custom PCB, all the surface-mounted components will also arrive from DigiKey. Once both the custom PCB and all its components have been received, the authors will move forward to assemble all the components onto the first custom PCB, which is explained in Section 8.2.

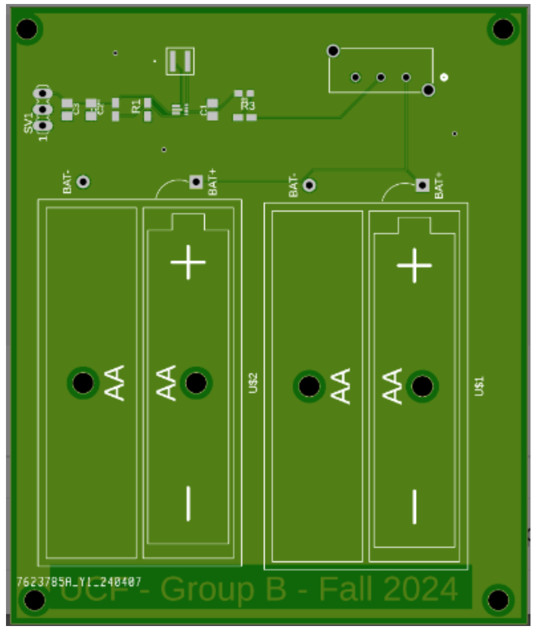


Figure 8.1.4 *Manufactured Custom PCB (First Revision)*

The Gerber file for manufacturing, with a build time of two days, comprised base material FR-4 with two layers, featuring dimensions of 81.3mm x 97.5mm and a PCB quantity of five. Classified as industrial/consumer electronics, it represented a single design delivered in a single PCB format, with a thickness of 1.6mm and green coloring. The material type was FR4-Standard TG 135-140, with a lead-free HASL surface finish and white silkscreen.

The design specified no impedance control and tented via covering, with no deburring/edge rounding. Fully tested via flying probe test, the custom PCB did not feature castellated holes or edge plating. Appearance quality adhered to IPC Class 2 Standard, with a confirmation production file and ink-jet/screen printing silk screen technology. The custom PCB maintained a board outline tolerance of ±0.2mm (Regular). Moreover, the design details and specifications aligned with JLCPCB's manufacturing requirements and standards.

As noted in Section 9.1.3, a second revision of a PCB was necessary and will resemble/coincide with the recommended layout depicted in the regulator's datasheet, which can be seen in Figure E.2. [26]

The second revision of the custom PCB was also designed using AutoDesk Fusion360. The following custom PCB includes explicit grounding and polygon outlining to distinguish a secure connection between the top and bottom layers.

The second iteration of the custom PCB without the top and bottom layer pour is listed below in Figure 8.1.5 with via's visible and strategically placed in accordance with Figure E.2. These vias continue to demonstrate the fact that the top and bottom board layers are connected securely. Additionally, the air wires were routed in a way that no wires overlapped, hence the red-colored wires. The blue-colored wires represent the few connections routed on the bottom layer to avoid overlap.

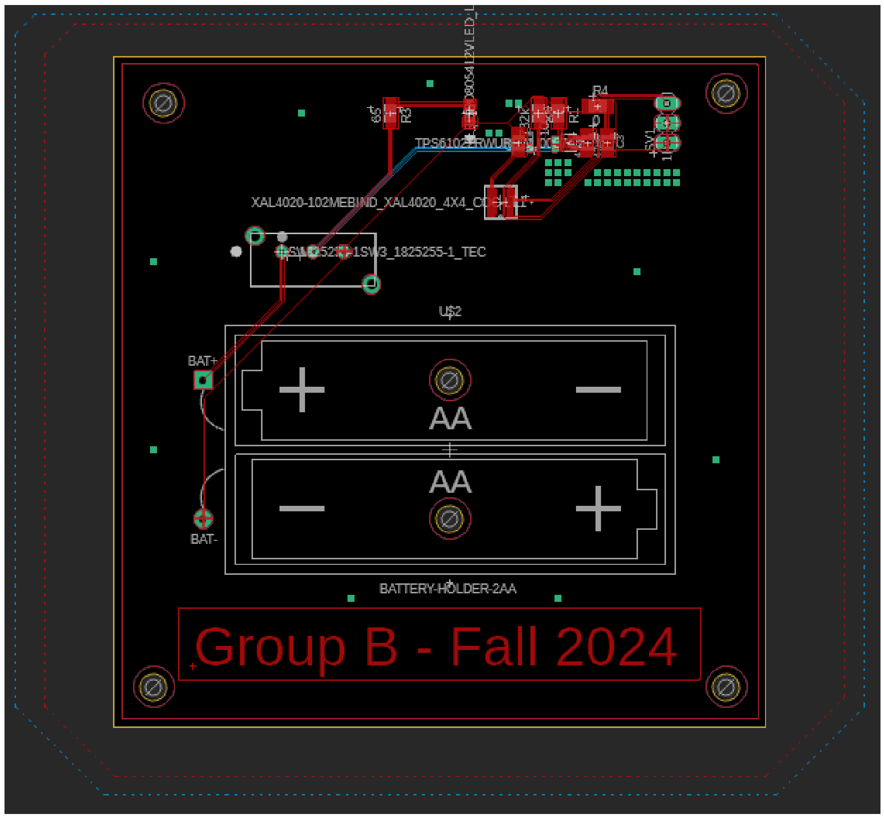


Figure 8.1.5 *Custom PCB Without Pour (Second Revision)*

The second iteration of the custom PCB with both the top and bottom layer pour is shown below in Figure 8.1.6.

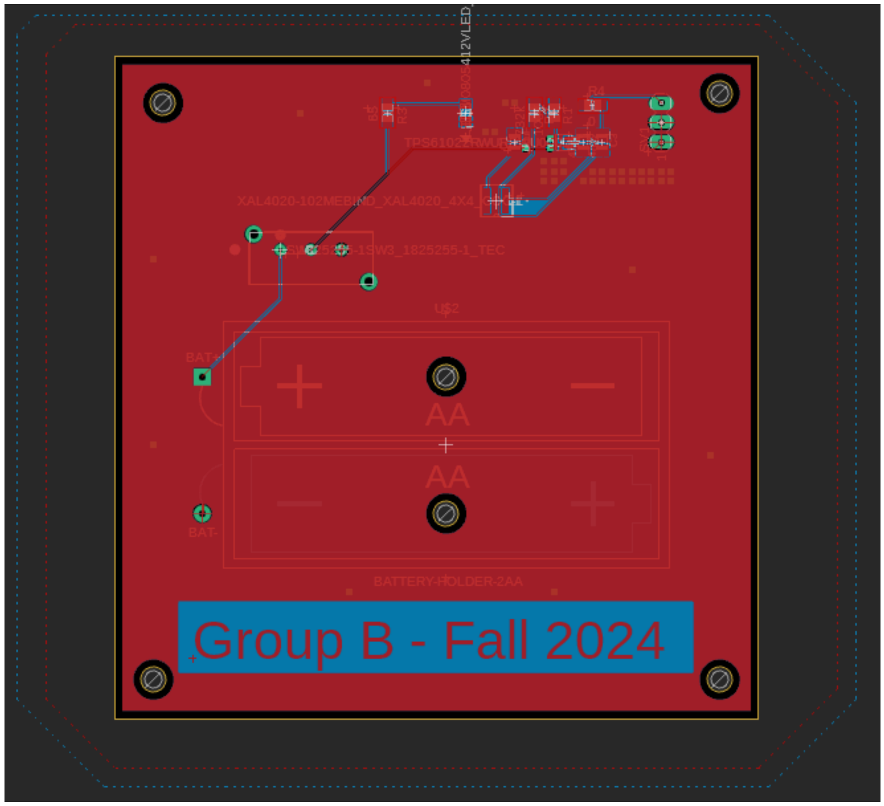


Figure 8.1.6 *Custom PCB With Pour (Second Revision)*

The second iteration of the custom PCB in its 3D version is listed below in Figure 8.1.7. The 3D model of the custom PCB houses the voltage regulator, closely resembling that in Figure E.2.

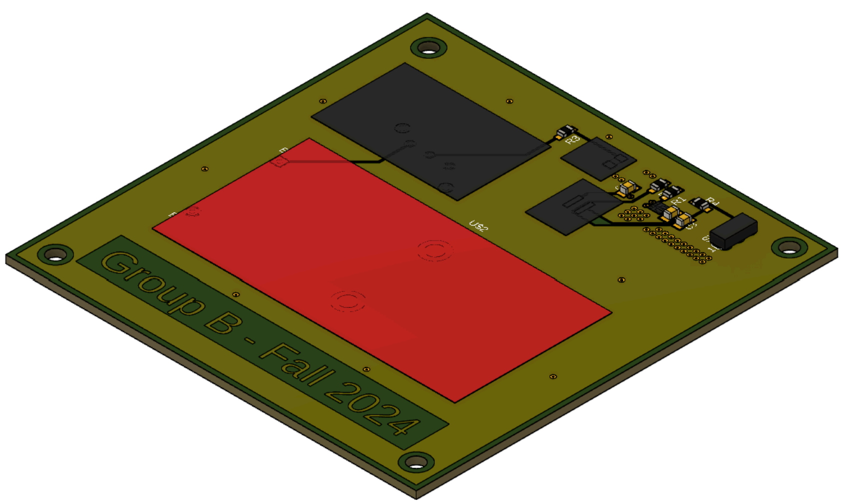


Figure 8.1.7 *3D Custom PCB (Second Revision)*

With the accommodations within the second iteration of the custom PCB, the authors expect a more seamless testing experience. Given that only one battery holder will be occupied, the expected voltage will be constant at every measurement using a digital multimeter (DMM).

8.2 Custom PCB Assembly

Once all the components of the custom PCB, along with the stencil, the authors performed a thorough manual assembly of the custom PCB. To begin the soldering process, the authors will position the stencil over the board, ensuring that the holes in the stencil mask align accurately with the SMD (Surface Mount Device) pads, which are the tiny areas of electrical conductors used to attach the electrical parts to the PCB.

Figure 8.2.1 *Custom PCB Stencil (First Revision)*

The authors then secured the stencil in place with double-sided tape on a lab work bench under a microscope connected to a monitor to ensure precise visibility. Then, using a solder paste syringe, the authors applied a small amount of solder paste onto the board just above the line of holes intended for solder application. The authors avoided directly inserting solder paste into the holes to prevent uneven distribution.

Next, using a flat plastic card, the authors spread the solder paste evenly across the stencil holes, applying gentle and consistent pressure to ensure thorough coverage without missing any areas of the pads. They verified the success of the solder paste application by ensuring that the SMD pads were fully covered with solder paste, with no visible metal showing through the stencil mask, which was purchased through JLCPCB.

Given that additional solder paste was required, the authors reused excess paste from the initial application. They did this while simultaneously being cautious not to over-apply it since that could directly lead to issues with component adhesion or even electrical connections. The authors repeated the process for each SMD component on the custom PCB. Once complete, they removed the double-sided tape and stencil and then cleaned the stencil thoroughly using isopropyl alcohol wipes to remove any residual solder paste.

The authors then placed the SMD components into their designated positions on the board using tweezers, ensuring they were centered and aligned properly without applying force. They then transferred the board to the Solder Reflow oven belt to reflow the SMD components.

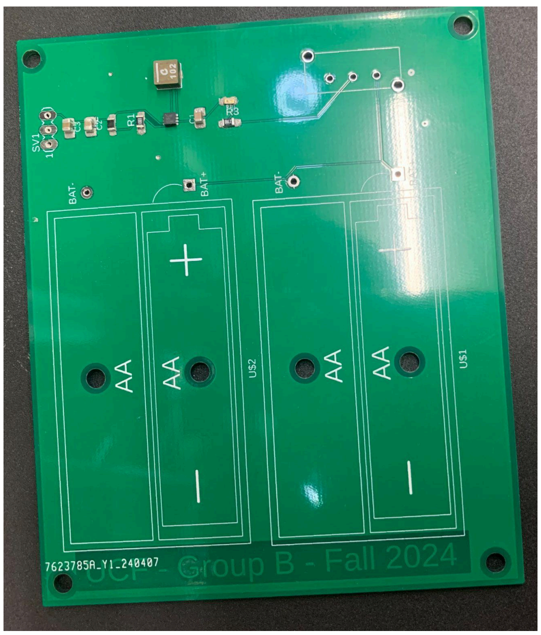
 **Figure 8.2.2** *Surface Mount Soldered PCB Top View (First Revision)*	 **Figure 8.2.3** Complete Soldered PCB Top View *(First Revision)*

Table 8.2 *Custom PCB Assembly (First Revision)*

Finally, after the reflow process was correctly completed, the authors hand-soldered the through-hole pin header and both of the AA battery holders onto the board. Manual soldering ensured precise and effective soldering for assembling the custom PCB's non-SMD components. Unfortunately, the mounting holes of the battery holders did not align with those of the footprint directly on the board. As a remedy, the authors used double-sided tape to secure the battery holders.

Once all the components were mounted properly, the authors used a Digital Multimeter to measure the output voltage of the PCB’s 3 pin header to ensure a functional PSU and to then shortly after begin the debugging development process. Further PCB testing is explained in Section 9.1.3. Given that a second revision of the custom PCB is necessary, the same process described in this section will be utilized identically when a new custom PCB and all its corresponding parts are ordered.

Chapter 9 System Testing and Integration

9.1 Early Software Testing

Currently, the authors have developed early-stage image calibration and image generation prototypes. On a Raspberry Pi 4B model, they developed a basic image calibration implementation using OpenCV in Python. This was an imperative step the authors took since the ultimate goal was to use a Raspberry Pi 4B for the final version of GPMS. This initial prototype allows the user to select the corners that the projector is able to reach, effectively calibrating the image to be superimposed over what the camera is viewing. From there, Canny Edge detection is done on the camera view and displayed through the projector, now highlighting the edges and vertices seen via the edge detection algorithm. This process is more thoroughly described in Section 9.2, which discusses integration. However, the frame still needs to be adjusted for barrel distortion in the camera lens and automated with a better visualization of the process.

Next, concerning image generation, the authors have developed an initial pipeline utilizing ControlNet + Stable Diffusion XL with the open-source GUI Automatic 1111. [5] [7] The table below, in the leftmost cell, is an example of an original existing image. This particular existing image is used due to its explicit edges and cartoonish style. In the middle cell of the table below is an example of an image generated using this new pipeline, using the prompt "*Cat eating a donut.*" Finally, on the rightmost cell of the table below is an animated GIF showing the tightness and accuracy of the edges in the original photo from the leftmost cell of the table. This accuracy will be vital in aligning the images with existing features.

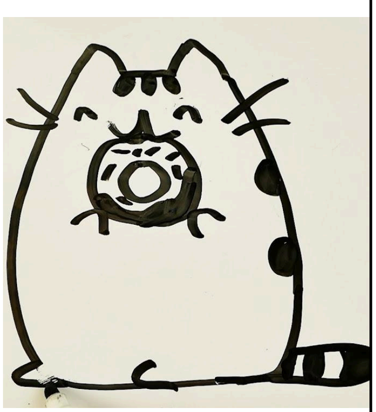
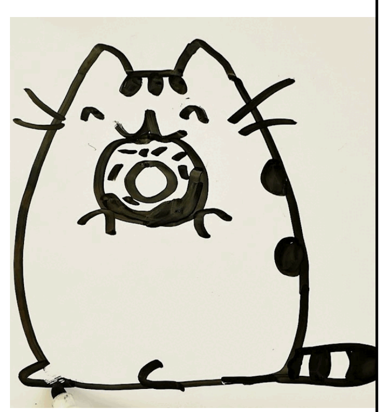
		

Figure 9.1.1	**Figure 9.1.2**	**Figure 9.1.3**

Original Image [51]	Image Generated w/ SDXL	Animated GIF to Display Similarities

Table 9.1 *Reprinted from "How to Draw Pusheen with Donut"* [51]

This process needs further exploration to enable ease of use for the final user. Much of this will come from prompt engineering algorithms, whether it be a Large Language Model (LLM) or a custom database that the authors can design to find keywords in user prompts and expand on what they believe the user is seeking to generate in their desired image. Finally, work should also be done to detect poorly generated images and regenerate them before they are returned to the user, another critical step in automating the process.

9.2 Integration Testing

In terms of integration, the authors have worked to put together a number of aspects that are key to representing the project as a whole. So far, this integration includes connecting many of the physical aspects of the GPMS to prototype how the final design will operate, specifically the projector, touch screen, Raspberry Pi, and Camera, as shown below.

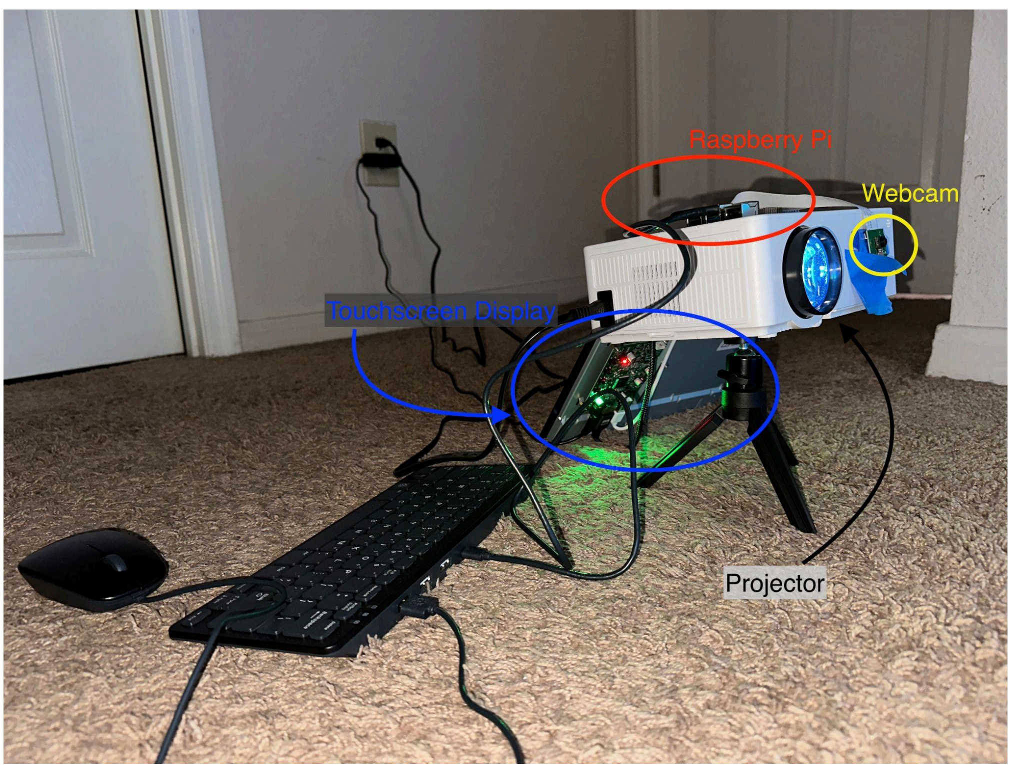


Figure 9.2.1 *Annotated Integration Image*

Once connected and set up, the authors loaded their Python code onto the Raspberry Pi, and after extensive debugging, tuning, and development, the application was started. During this time, a whiteboard was placed as a backdrop for the projected scene, and images/words were drawn.

The prototype software, as seen in Appendix D, was built using Python to control OpenCV, which worked well for this initial prototype; however, going forward, the Qt framework will be critical to managing windows and displays better, as this initial prototype has an initialization sequence where the user must drag one window onto the other display before continuing as to circumvent a number of issues with OpenCV integration on Pi OS.

Once initialized, though, a white screen is shown on the projector to allow for better calibration by the user with a camera stream from the front webcam displayed on the touchscreen. In this “reset” stage, the user selects the outer four corners of the projected scene, as shown below.

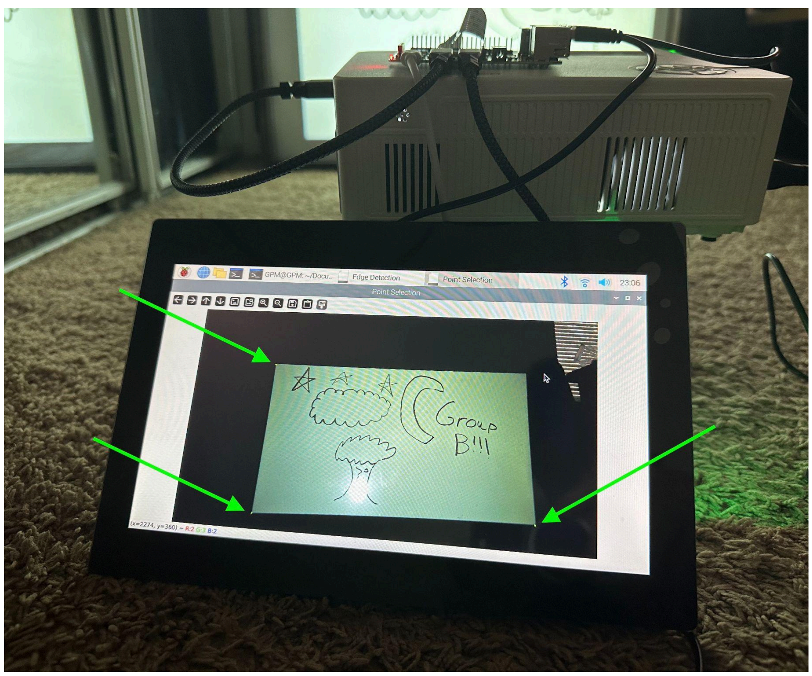


Figure 9.2.2 *Prototype Image Calibration*

Once selected, the device captures a still frame of the scene using the front webcam, performs edge detection on it, and traces the outline of the scene, as shown below.

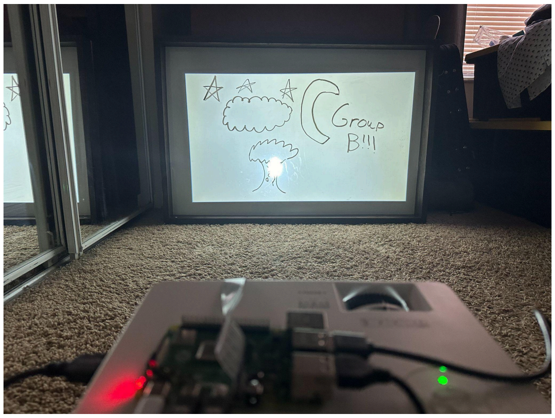


Figure 9.2.3 *Initial Drawing*

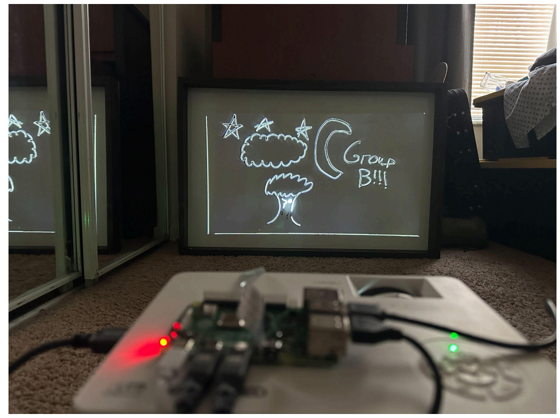


Figure 9.2.4 *Projected Outline*

Table 9.2 *Prototype Image Alignment Performance*

From here, the threshold can be adjusted using ‘O’/’P’ and ‘K’/’L’ to decrease and increase the low and high thresholds, respectively. Additionally, the user can press ‘R’ to reset the 4 edge points and ‘U’ to update the threshold. When updating the projection, the projector flashes white for 3 seconds to allow time for the webcam exposure to compensate before grabbing a still frame to perform the edge detection algorithm. Finally, this image is displayed again as before.
This prototype effectively proves the core process for aligning the edges of the projection against surface features. However, this implementation also provides valuable insight into key details for future versions. For example, the fine white edges on the left and bottom edges of the projection are the result of not being able to precisely select the corner due to the size of the image on the screen. Because of this, the edge detection algorithm detects the edge of the projection screen instead of actual features. Therefore, going forward, a better plan would be to allow the user to zoom in on the image to select the corners more precisely. Additionally, the bright shine of the projector on the semi-reflective whiteboard made detecting the black marker difficult, as seen in the bottom middle of the above images, not totally highlighting the marks in the shine. This can be easily solved by using an additional camera such that neither glare is in the same spot, allowing software to effectively overwrite the issue and create a complete image without glare.

9.3 PCB Testing / Evaluation

The authors tested the output voltage of their custom PCB to ensure precision and quality in their Battery Management System. They prepared a power bench, setting the output to the approximate voltage of the batteries and adjusting the current to a safe level to protect the components. They then connected these leads to the pads on the PCB. Using a Digital Multimeter, they checked the voltage output at the 3 pin header, confirming consistent 5V levels indicative of PCB integrity.

Following initial tests, they connected the batteries, monitoring the output's voltage stability against their expectations. Successful matching of voltage levels confirmed the PCB's performance.

Testing the 5V output is crucial for the Raspberry Pi 4B, which relies on a stable 5V supply for optimal functionality. This voltage powers the Raspberry Pi's core components like the processor and peripherals, ensuring compatibility and preventing voltage-related issues (i.e. over or under voltage). The testing also verifies the PCB's reliability as a power source, maintaining consistent voltage for stable Raspberry Pi operation and safeguarding against system failures.

In conclusion, testing the PCB's output voltage is essential for ensuring the Raspberry Pi 4B's compatibility, reliability, and safe operation within portable applications, enhancing its performance and longevity.

Figure 9.3.1 *PCB With Battery Supply*

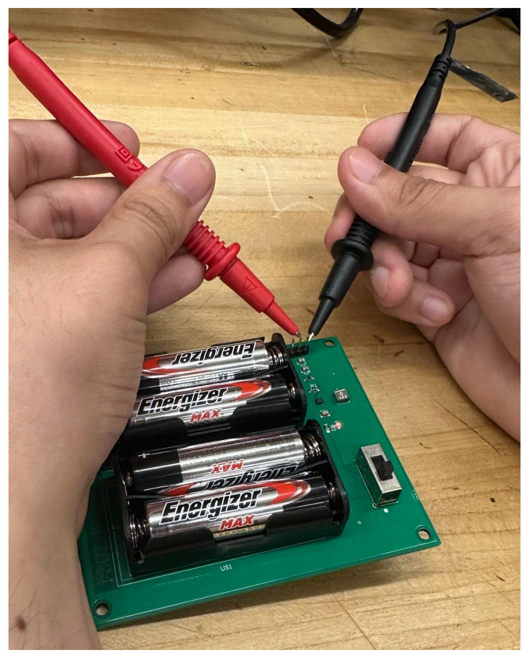


Figure 9.3.2 *PCB Testing using DMM*

Table 9.3.1 *PCB Testing*

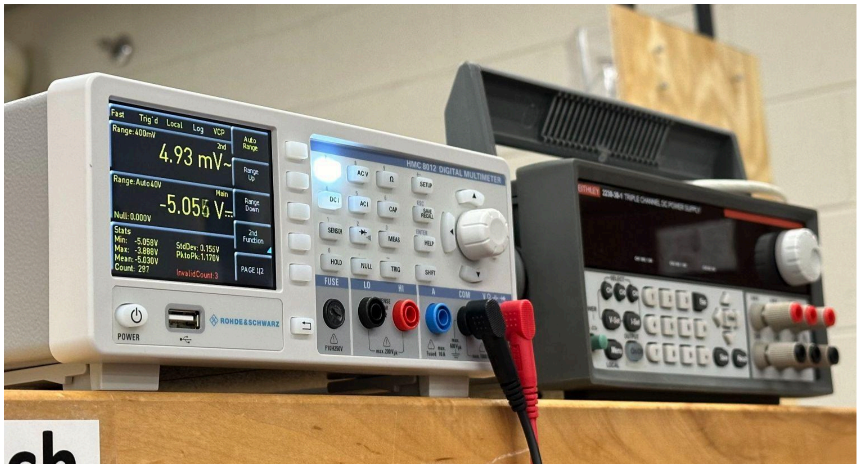


Figure 9.3.3 *DMM Voltage Reading*

After testing the PCB using a DMM, the reading was ~5V (ignoring the negative sign shown in Figure 9.3.2), as expected for what the Raspberry Pi 4B requires, and provides tolerance for a maximum of 6V which is not to be exceeded. [58] Therefore, the PCB was functioning properly.
However, the authors turned the on/off switch a couple of times and kept measuring the output voltage to ensure that 5V was being read each time by the DMM. Additionally, the authors used a breadboard to measure current, and the DMM read no current flow. After continuous testing, the authors fried the regulator and therefore the V_{in} would not be valid anymore for the regulator.

Therefore, with continuous testing, the DMM read 1.5V from the 3 pin header output instead of the 5V before. As a result the authors took steps of action to begin the revision process of the PCB. In Figure 6.2.2, a new schematic for the PCB was created and temporarily implemented (with the assistance of Dr. Weeks) on the current non-functioning PCB using a physical wire to route the LED negative terminal to ground. In the temporary physical implementation, the connection from the LED to the capacitor was shortened by splicing the connection which directly allows for a proper input voltage delivery. The temporary implementation can be seen in Table 9.3.2.

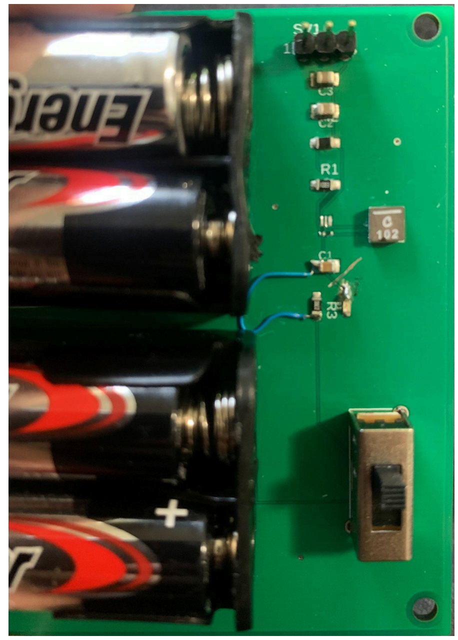
 **Figure 9.3.4** Corrected *PCB With Battery Supply*	 **Figure 9.3.5** *Up Close Corrected PCB With Battery Supply*

Table 9.3.2 *PCB Modifications (First Revision)*

9.4 Plan for Senior Design 2

While more explicit steps and plans are elucidated in Chapter 10's sections on "Senior Design 2 Milestones", 10.2, and "Technical Milestones," 10.5, this section will briefly discuss the higher level objectives in software and hardware integration for Senior Design 2, considering where the authors are in their current stage.

In terms of planned Software Development, the general outline is as follows:

- **User Interface (UI) Development**: Design a user-friendly interface that allows easy interaction with the system and aligns with the overall design plans discussed in Section 7.3. The UI will enable users to seamlessly enter custom prompts and view generated images, providing an intuitive platform for both novice and experienced users.
- **Networking Infrastructure**: To facilitate robust communication between the user interface and server, we will develop a networking infrastructure as discussed in Section 7.2.4. This setup will handle requests and responses efficiently, ensuring smooth data flow and system responsiveness.
- **Python Server Development**: A dedicated Python server will be created to manage the generative design stack, which will reside on the server as

discussed in Section 7.4. This server will handle requests and handle the automation of image generation as it is discussed in the next objective.

- **Automation of Image Generation**: In the initial prototype, each image was generated with the authors supervision, so automating the image generation process is essential. By streamlining operations from prompt input to image output, the aim is to minimize user wait times and optimize system performance.
- **Exploration of Prompt Engineering**: To enhance the relevance and quality of generated images, prompt engineering techniques need to be further explored with respect to the automation methods previously discussed. These methods will analyze user inputs to extract and expand upon key themes, potentially integrating a Large Language Model (LLM) to refine the system's interpretative capabilities.

With respect to Hardware plans:

- **Battery System Iterations**: The authors will continue to refine the battery system to ensure longer operational life and higher efficiency. This involves evaluating current performance, identifying bottlenecks, and implementing improvements.
- **Physical Enclosure Design**: A robust physical enclosure will be designed to house the projector, camera, and touchscreen monitor setup. This enclosure will protect the hardware from environmental factors and facilitate ease of transport and setup.
- **Sensor Vision Stack Development**: Further development of the sensor vision stack is planned. By enhancing the capabilities of these sensors, the aim is to improve the accuracy and quality of image calibration and alignment, which is crucial for precise image overlay on varied surfaces.

Chapter 10 Administrative Content

10.1 Budget and Financing

The prices provided are based on online research and manufacturer quotes and are, therefore, estimates. These costs are subject to change upon implementation of the final product.

Currently, the project has no corporate sponsors, and the group will cover the majority of associated expenses aside from those specifically mentioned below. Should the authors attain further sponsors or donations, they will immediately declare that in the overall Bill of Materials (BOM) in Table 10.1.

ITEM	QUANTITY	PRICE ESTIMATE (USD)
Custom PCB & Stencil	5 & 1	30.42

Custom PCB Components	13	34.84
Raspberry Pi 4B	1	60.79
Camera	1	5.00
Monitor	1	89.99
Battery Pack	24	15.00
Projector	1	63.89
AMD Machine*	1	3244.44
TOTAL (ESTIMATE)	**N/A**	**~ 3,544.37**
AFTER DONATIONS	**N/A**	**~ 299.93**

Table 10.1 *Overall Project BOM*

*denotes a component lent from AMD.

All the authors agreed to evenly split the total cost (after donations) upon completing Senior Design 2 (EEL 4915L).

10.2 Project Milestones

The dates outlined in the table provided serve as crucial milestones for the group, marking hard deadlines that all the authors must adhere to in order to manage the progress effectively. The authors have established a weekly meeting schedule to address individual concerns and collectively ensure that they maintain momentum in advancing the overall design objectives.

During the initial phase of Senior Design 1, the primary focus will be on fulfilling administrative responsibilities. This includes meeting deadlines for outlining all project details and specifications meticulously. Once these foundational tasks are successfully accomplished, the authors will shift their focus entirely toward the development phase of the project. This concentrated effort is aimed at ensuring substantial progress by the end of Senior Design 1, setting a solid foundation for the subsequent phase, Senior Design 2.

10.2.1 Senior Design 1

The authors expect the majority of all the project details to be completed and finalized by the end of Senior Design 1 (EEL 4914C) with the expectation of not making any new or further adjustments beyond this phase of the project of project planning for GPMS. Below is a table that comprehensively outlines the entire roadmap of Senior Design 1. This table strategically breaks down the roadmap into stages with respective milestones and hard deadlines.

Stage	Milestone	Deadline
1	Recruit Group Members	1/11/24
2	Individual/Collaborative Research	1/15/24
3	Reach Project Decision	1/24/24
4	Submission of Divide & Conquer Document	2/2/24
5	Divide & Conquer Feedback Meeting	2/5/24
6	Completion of 45 Page Document	3/10/24
7	Finalize BOM	3/15/24
8	Completion of User Interface	3/20/24
9	Submission of 45 Page Document	3/29/24
10	45 Page Feedback Meeting	3/30/24
11	Order First Custom PCB	4/5/24
12	Test First Custom PCB	4/10/24
13	Begin Implementing CV Algorithms	4/15/24
14	Submission of 90 Page Document	4/19/24
15	SD1 Reviewer Committee Check In	4/22/24

Table 10.2.1 *Senior Design 1 Timeline*

10.2.2 Senior Design 2

With the majority of all the project details set to be completed by the end of Senior Design 1, the authors plan to solely use the allotted time for Senior Design 2 (EEL 4915L) to prototype, test, and debug GPMS in order to complete the project well before the final deadline.

Below lies a table that comprehensively outlines the entire roadmap of Senior Design 2 broken down into stages with respective milestones and hard deadlines.

Stage	**Milestone**	**Deadline**
1	Assemble Second Custom PCB	5/16/24
2	Test Second Custom PCB	5/16/24
3	Achieve Working Custom PCB	5/16/24
4	Design Hardware Enclosure	8/23/24
5	Overall CV Software Stack	10/31/24
6	Finishing Touches to UI Design	11/20/24
7	Test Entire Project	11/20/24
8	Reviewer Committee Demonstration	12/1/24

Table 10.2.2 *Senior Design 2 Timeline*

10.3 Table of Work Distributions

Effective work distribution is crucial for the success of the development of GPMS, ensuring that tasks are completed efficiently and on schedule. In the authors' project, which involves various aspects such as PCB design/hardware integration, computer vision software stack, UI interface design, prototype design, generative AI pipeline, and overall design report, assigning responsibilities to the appropriate team members is paramount. For instance, Francisco Soriano takes the lead in PCB design/hardware integration, with support from Declan Carter and Victoria Moreno. This division of labor ensures that technical expertise is utilized effectively, with primary responsibility resting on the team member best suited for the task.

Similarly, in computer vision software stack development, Declan Carter spearheads the efforts, supported by Victoria Moreno and Francisco Soriano. This arrangement ensures a focused approach to software development, leveraging Declan's proficiency in the field while benefiting from the contributions of other team members. Furthermore, in tasks such as UI interface design and prototype development, Victoria Moreno will take the lead, with support from the other two team members.

This distribution of responsibilities allows each team member to play to their strengths, maximizing productivity and expertise in their designated areas. Ultimately, by carefully assigning tasks to primary and secondary persons based on their skills and capabilities, the authors ensure that efficient progress across all facets of the project, culminating in a successful outcome.

The table below visualizes a distribution of work tasks between all the authors with respective primary and secondary persons.

Stage	**Component Owner**	**Secondary Owners**
PCB Design/Hardware Integration	Francisco Soriano	Declan Carter/Victoria Moreno
Computer Vision Software Stack	Declan Carter	Victoria Moreno/Francisco Soriano
UI Interface Design	Victoria Moreno	Francisco Soriano/Declan Carter
Prototype Design	Declan Carter	Victoria Moreno/Francisco Soriano
Generative AI Pipeline	Victoria Moreno	Declan Carter/Francisco Soriano
Design Report	Francisco Soriano	Declan Carter/Victoria Moreno

Table 10.3 *Work Distribution*

Note: Beyond the fundamental stages necessary to the development and completion of GPMS, the website, which contains all documents related to the project, must be maintained. Francisco Soriano, in addition to creating the website in HTML and CSS, will take on the responsibility to maintain the website throughout the duration of both Senior Design 1 and 2.

10.4 Project Management Strategy

Given that Jira is a widely used project management software in industry, primarily designed for agile teams, the authors decided upon using this software throughout the duration of Senior Design 1 and 2. Additionally, since Jira is scalable to work for small groups to large groups, the authors determined that their three person team is sufficient to benefit from all that Jira has to offer and benefit. Jira's usefulness to the team stems from several stand out features regarding the following information listed in the table below. [59]

Stand Out Features	Description
Task Tracking	Jira allows the authors to create and track tasks, issues, bugs, and user stories throughout the project lifecycle. This is extremely helpful as it organizes work and ensures that no task is left unaddressed.
Agile Methodology Support	Jira provides visual boards for tracking progress to manage projects effectively.
Collaboration	Jira promotes collaboration among team members by providing features such as commenting, mentioning, and assigning tasks, which in turn fosters communication and transparency amongst the authors.
Reporting and Insights	Jira offers reporting and dashboard features that provide the authors insight into project progress, team performance, and bottlenecks to identify significant areas for improvement.
Remote Work Support	Jira's cloud-based version has allowed the authors to access and manage their tasks from anywhere on or off campus. It has allowed the fostering of collaboration and productivity regardless of physical location.

Table 10.4 *Project Management Strategy*

Jira's versatility, flexibility, and robust feature set make it an invaluable tool for project management for the authors due to several key, stand out, reasons. Moreover, Jira fosters collaboration among team members by offering features such as commenting, mentioning, and task assignment, enhancing communication and transparency within the team and among the authors.

Furthermore, Jira's cloud-based version enables remote work, allowing the authors to access and manage tasks from anywhere, whether on or off campus. This flexibility not only enhances collaboration but also boosts productivity, regardless of the authors' physical locations. Overall, Jira's comprehensive suite of features makes it an indispensable asset for authors to manage the project efficiently.

10.5 Technical Milestones

The Gantt Chart in the image below, provided through Jira Management Software, is a high-level monthly overview of the various technical milestones stretching from January to September in quarter intervals.

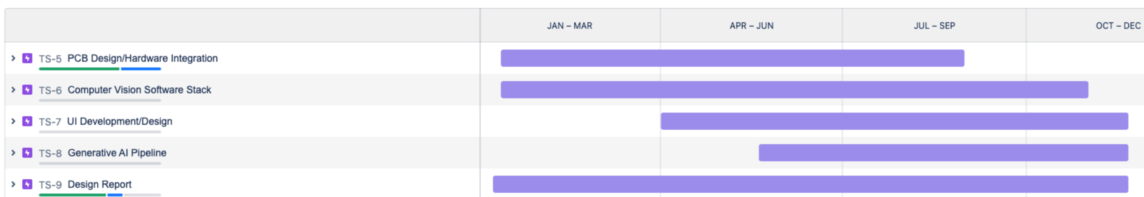


Figure 10.5 *Jira Overview*

The following list depicts a high-level view of what significant, overarching tasks/stages are needed to contribute to the successful completion of GPMS.

The main overarching tasks/stages are listed as follows and will be explicitly listed in their corresponding tables in the following sections:

- PCB Design/Hardware Integration
- Computer Vision Software Task
- UI Development/Design
- Generative AI Pipeline
- Design Report

10.5.1 PCB Design/Hardware Integration

Designing the custom PCB is imperative as it serves as the principal PSU. In fact, the initial custom PCB design serves as the kickstart of the overarching goal of the PCB Design/Hardware Integration stage of the development of GPMS.

In the project timeline, the process of ordering the PCB marks a critical milestone. This task involves coordinating with suppliers, specifying design requirements, and ensuring timely delivery to maintain project momentum, and undergoing revisions for the most optimal PSU. The timely completion of this step is essential as it lays the groundwork for subsequent phases, particularly in hardware integration. By adhering to this deadline, the authors ensure that the project stays on track and avoids potential delays that could very well negatively impact the overall timeline of the GPMS development process.

Following the PCB ordering process, the next significant task is the testing of the PCB to ensure a proper power supply to the Raspberry Pi 4B. Assuring a proper and seamless power supply to the Raspberry Pi 4B involves rigorous testing procedures to verify the functionality and reliability of the power supply system, crucial for the stability and performance of the Pi.

Meeting the proper deadlines for PCB testing is extremely imperative as it allows for any necessary adjustments or troubleshooting before advancing to subsequent stages of development. Additionally, timely testing ensures early detection of any potential issues, mitigating risks and facilitating smoother progress in the project's hardware integration phase. The table below illustrates an overview of the required tasks, regarding the PCB Design/Hardware Integration, and also the due dates that the authors intend to meet promptly.

Task	**Due Date**
Design First Custom PCB	March 25th
Order First Custom PCB	March 30th
Testing PCB to ensure proper power supply to Raspberry Pi 4B	April 17th
Design Second Custom PCB	April 18th
Order Second Custom PCB	April 20th
Test Second PCB Iteration	May 13th
Final Functioning PCB	August 30th

Table 10.5.1 *PCB Design/Hardware Integration Deadlines*

10.5.2 Computer Vision (CV) Software Stack

The completion of the overall computer vision (CV) software stack represents a critical milestone in the authors' project, as it signifies the culmination of various development stages aimed at enhancing image processing capabilities. Beginning with the creation of prototype software to test edge distortion, followed by a similar prototype for lens distortion, the team is currently systematically addressing key aspects of image correction and enhancement. Subsequent tasks, such as aligning images manually over the subject and performing edge detection with user prompts for alignment before the specific set deadline, demonstrate the team commitment to refining the software's functionality in detecting and correcting distortions effectively.

As the authors' progress, the timeline includes experimentation with additional sensors such as multiple cameras, and exploring automated methods for image alignment. These tasks underscore their dedication to innovation and continuous improvement, as they strive to optimize the software's performance and versatility.

Finally, the creation of a barebones user interface (UI) to guide users through the image correction process not only enhances usability but also marks a significant step towards the practical application of the CV software stack. Through meticulous planning and execution of these tasks, the authors' aim to deliver a robust and user-friendly solution that meets the project's objectives effectively. The table below illustrates an overview of the required tasks, regarding the Computer Vision Software Stack, and also the due dates that the authors intend to meet promptly.

Task	**Due Date**
Create Prototype Software to Test Edge Detection	February 15th
Align Images Manually Over Subject	April 15th
Perform Edge Detection & Prompt User to Align Edges	April 17th
Create Prototype Software to Test Lens Distortion	April 22nd
Experiment w/ Additional Sensors (Multiple Cameras)	August 15th
Experiment w/ Automated Methods	August 30th
Overall CV Software Stack	October 31st

Table 10.5.2 *Computer Vision Software Stack Deadlines*

10.5.3 UI Development/Design

The task of utilizing Figma Interface Design Tools to prototype UI design for future development represents a pivotal step in the authors' project's user interface (UI) development process. Figma's versatile and collaborative platform offers an ideal environment for them to conceptualize, iterate, and refine the UI design, ensuring that it meets the project's requirements and user expectations. By leveraging all of Figma's intuitive features and extensive library of design elements, they can create interactive prototypes that provide a realistic preview of the final product, facilitating valuable feedback and validation from stakeholders.

Meeting the deadline for UI prototyping with Figma not only demonstrates the author’s commitment to timely project execution but also sets the stage for seamless integration with subsequent development phases. The prototypes generated through Figma serve as invaluable assets for guiding the development team in implementing the UI design accurately and efficiently. Furthermore, Figma's cloud-based platform enables seamless collaboration and version control, allowing team members to work together effectively regardless of their location.

By harnessing the power of Figma, the authors aim to deliver a user-friendly and visually appealing UI that enhances the overall usability and experience of the project. Moreover, once the conceptual Figma UI is finalized, the actual implementation of it will be seamless given that there will be absolutely no time wasted on making adjustments to the UI. The table below illustrates an overview of the required tasks, regarding the UI Development/Design and the due dates that the authors intend to meet promptly.

Task	**Due Date**
Using Figma Interface Design Tools to Prototype UI Design	April 1st
Create Barebones UI to Walk User Through the Process	September 15th
Integrate CV Stack	September 30th
Finishing Touches to UI Design	November 20th

Table 10.5.3 *UI Development/Design Deadlines*

10.5.4 Generative AI Pipeline

The completion of the overall generative AI pipeline is another significant milestone in the project, representing the culmination of various development stages aimed at harnessing the power of artificial intelligence for creative generation.

Beginning with the task of securely accessing AMD's server via SSH, the team establishes the foundational infrastructure necessary for subsequent phases of pipeline development. Testing generative AI UI platforms and experimenting with edge detection applications promptly, demonstrate the team's close commitment to exploring a diverse set of methodologies and technologies to enhance the pipeline's functionality and performance.

As the team progresses, tasks such as determining the pipeline structure using ControlNet, identifying areas for prompt expansion, and defining packet formatting, underscore the meticulous approach to pipeline design and implementation.

Through diligent execution of these tasks, the authors aim to deliver a robust and versatile generative AI pipeline that meets the project's objectives and exceeds expectations. The table below illustrates an overview of the required tasks regarding the Generative AI Pipeline and the due dates that the authors intend to meet promptly.

Task	**Due Date**
Test Generative AI UI Platforms	February 20th
SSH into AMD's Server	May 30th
Determine Packet Formatting	August 1st
Test Raspberry Pi Connection to Generative Stack	August 15th
Determine Pipeline/Edge Detection using ControlNets	August 22th
Determine Use of Prompt Expansion	September 15th
Overall Generative AI Pipeline	November 20th

Table 10.5.4 *Generative AI Pipeline Deadlines*

10.6 Design Report

The task of producing a comprehensive 45-page report by the deadline is a pivotal point in the overall project timeline of Senior Design 1, requiring meticulous documentation and analysis of the findings, methodologies, and outcomes. The halfway report represents a crucial opportunity to convey the project's progress, insights, and challenges thus far, providing stakeholders with a comprehensive understanding of the work to date. By meeting the halfway document deadline, the authors demonstrate their commitment to transparency and accountability, ensuring that the project remains on track and aligned with its objectives.

Following the completion of the 45-page report, the focus shifts towards creating a finalized 90-page report due. This extended document serves as the culmination of the authors' efforts, offering an in-depth exploration of the project's scope, methodologies, results, and conclusions. By expanding upon the initial report, the authors provide a more comprehensive analysis of their findings, incorporating additional details, discussions, and recommendations to offer stakeholders a deeper insight into the project's significance and implications.

Through the timely delivery of both reports, the authors aim to showcase the rigor and dedication with which the authors approach their work, ensuring that the project's outcomes are effectively communicated and understood by all relevant parties. Jira has allowed the authors to set technical deadlines that keep the authors well ahead and also on schedule to complete GPMS, leaving ample room for debugging and potential modifications for the duration of the project.

Chapter 11 Conclusion

Throughout the duration of Senior Design 1, the authors have learned first hand the importance of planning strategically and collectively to design a Generative Projection Mapping System, GPMS, capable of immersing all audiences through projection mapping.

The authors are ecstatic to continue on an ambitious project aimed at creating GPMS, which seeks to revolutionize storytelling environments through dynamic image projection onto various structures. Their primary goal is to continue utilizing advanced computer vision techniques to intelligently detect and project images onto unique features such as pillars, windows, or drawings. GPMS, with a vital PSU, integrates a responsive touchscreen monitor to allow users to input text or prompts, initiating the generation of images that seamlessly align with the structure's features.

In their pursuit of enhancing user customization, the authors aim to implement sensitivity and threshold controls to increase user control over image representation which empower all users to intricately adjust the outline of the structure according to their preferences, ensuring a personalized experience.

Furthermore, the authors plan to enable users to highlight specific structural features through an intuitive interface, allowing for a personalized visual experience. They also aim to streamline the user experience by integrating an automatic calibration process into the device, minimizing user input and ensuring optimal image display without unnecessary complexity. The authors plan to continue meeting weekly for the remainder of the GPMS development journey.

Finally, the authors aim to meet the set deadlines to ensure ample time in accomplishing the project adequately. In conclusion, the authors' project aims to push the boundaries of immersive storytelling environments while providing users with unprecedented control and customization options throughout the duration of Senior Design 2.

Appendix

Appendix A - Pledge of Originality

We, the members of Senior Design Group B of the University of Central Florida, hereby pledge and affirm that our final report, submitted in partial fulfillment of the requirements for the EEL 4914C and EEL 4915L courses, represents our own original work.

All ideas and findings presented herein are the result of our collective effort and intellectual endeavor. Where we have consulted the work of others, it is duly acknowledged within the body and Appendix-B of the report, ensuring that credit is given where it is due. This document is a product of our collaboration as a group. Each team member has contributed significantly to the conceptualization, research, analysis, and writing processes. We have openly shared knowledge and worked together to overcome challenges, ensuring a comprehensive and cohesive final product. We have adhered strictly to the ethical guidelines provided by our Faculty Mentors and the University of Central Florida's academic honor code. We have made every effort to document our methodologies, data sources, and analysis techniques transparently, allowing for the reproducibility and scrutiny of our work.

By signing this pledge, we commit to the truthfulness of the above statements.

[Signed]

Francisco Soriano Declan Carter Victoria Moreno

Appendix B - References

[1] OpenAI. "ChatGPT." Chat.openai.com, OpenAI, 30 Nov. 2022, chat.openai.com/.

[2] EVERY Disney World and Disneyland Castle TRANSFORMATION in HISTORY! | the Disney Food Blog. 26 May 2020, www.disneyfoodblog.com/2020/05/26/every-disney-world-and-disneyland-castle-transformation-in-history/. Accessed 30 Jan. 2024.

[3] Sylt, Christian. "The Magic behind the Scenes of Universal's Harry Potter Show." Forbes, www.forbes.com/sites/csylt/2020/03/07/the-magic-behind-the-scenes-of-universals-harry-potter-show/. Accessed 30 Jan. 2024.

[4] Sandvik, A., & Zhang, J. (2023). Dynamic Projection Mapping A Senior Project presented to. https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1377&context=cpesp

[5] Luma Box. (n.d.). Projection Mapping. Luma Box. https://www.lumabox.com/

[6] Chauhan, V. K., Zhou, J., Lu, P., Molaei, S., & Clifton, D. A. (2023, August 10). *A brief review of Hypernetworks in deep learning*. arXiv.org. https://arxiv.org/abs/2306.06955#:~:text=Hypernetworks%2C%20or%20hypernets%20in%20short,sharing%2C%20and%20model%20compression%20etc.
[7] Zhang, L., Rao, A., & Agrawala, M. (2023, November 26). *Adding conditional control to text-to-image diffusion models*. arXiv.org. https://arxiv.org/abs/2302.05543
https://github.com/lllyasviel/ControlNet
[8] Chauhan, V. K., Zhou, J., Lu, P., Molaei, S., & Clifton, D. A. (2023, August 10). *A brief review of Hypernetworks in deep learning*. arXiv.org. https://arxiv.org/abs/2306.06955#:~:text=Hypernetworks%2C%20or%20hypernets%20in%20short,sharing%2C%20and%20model%20compression%20etc.

[9] MCL_Admin. (2021, April 6). FPGA vs. Microcontroller | What is the Difference? Mcl. https://www.mclpcb.com/blog/fpga-vs-microcontroller/
[10] Gregersen, C. (2023, September 6). *A Complete Guide to Development Boards*. Nabto. https://www.nabto.com/development-board-guide/
[11] *How to choose a microcontroller for your task: the pros and cons of various MCUs*. (2022, May 4). 2Smart. https://2smart.com/docs-resources/articles/how-to-choose-a-microcontroller-for-your-task-the-pros-and-cons-of-various-mcus
[12] Raspberry Pi LTD. (n.d.). Buy a Raspberry Pi 3 Model B. Raspberry Pi. https://www.raspberrypi.com/products/raspberry-pi-3-model-b/
[13] *Raspberry Pi 4 Computer Model B*. (2021). https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-product-brief.pdf
[14] ESP32-WROOM-32 Datasheet. (n.d.). https://www.espressif.com/sites/default/files/documentation/esp32-wroom-32_datasheet_en.pdf
[15] MSP-EXP430FR6989 | Buy TI Parts | TI.com. (n.d.). Www.ti.com. Retrieved February 17, 2024, from https://www.ti.com/product/MSP-EXP430FR6989/part-details/MSP-EXP430FR6989#:~:text=Features%20for%20the%20MSP-EXP430FR6989%201%20MSP%20ULP%20FRAM-based
[16] Amazon.com. (n.d.). Projector, Video Projector Multimedia Compatible Smartphone. Retrieved April 11, 2024, from https://www.amazon.com/Projector-Video-Projector-Multimedia-Compatible-Smartphone/dp/B07MTCMHZX/
[17] Amazon.com. (n.d.). Projector, Bluetooth 100Screen Compatible Smartphone. Retrieved April 11, 2024, from https://www.amazon.com/Projector-Bluetooth-100Screen-Compatible-Smartphone/dp/B0B28G5Y4R/
[18] Amazon.com. (n.d.). Bluetooth Projector Supported Portable Compatible. Retrieved April 11, 2024, from

https://www.amazon.com/Bluetooth-Projector-Supported-Portable-Compatible/dp/B0BGXKQ4MP/
[19] Microsoft. (2024). Microsoft Copilot: Your everyday AI companion. Copilot.microsoft.com. https://copilot.microsoft.com/
[20] Kumar, P. (2021, November 11). *MIPI camera vs USB camera – a detailed comparison*. E-Con Systems. https://www.e-consystems.com/blog/camera/technology/mipi-camera-vs-usb-camera-a-detailed-comparison
[21] Davis, C. (2023, July 21). *4 touch panel types - explained*. ViewSonic Library. https://www.viewsonic.com/library/business/touch-panel-types-explained/
[22] *Linear vs. Switching Regulators | Renesas*. (n.d.). Www.renesas.com. https://www.renesas.com/us/en/products/power-power-management/linear-vs-switching-regulators
[23] Simpson, C. (n.d.). *Linear and Switching Voltage Regulator Fundamental Part 1 National Semiconductor LINEAR VOLTAGE REGULATORS*. https://www.ti.com/lit/an/snva558/snva558.pdf
[24] *Using a Switching Regulator vs. Linear Regulator for DC-DC Conversion*. (n.d.). Resources.pcb.cadence.com. https://resources.pcb.cadence.com/blog/2020-using-a-switching-regulator-vs-linear-regulator-for-dc-dc-conversion
[25] Power Designer. (n.d.). Webench.ti.com. https://webench.ti.com/power-designer/switching-regulator/select
[26] TPS61022 8-A Boost Converter with 0.5-V Ultra-low Input Voltage. (n.d.). Retrieved February 11, 2024, from https://www.ti.com/lit/ds/symlink/tps61022.pdf?ts=1707679005182
[27] Dr. Samuel Richie, Dr. Avra Kundu (2023, September). EEL 3926L Junior Design Part 2 - WEBENCH Exercise [Lecture notes]. University of Central Florida.
[28] How to Choose Batteries | REI Expert Advice. (n.d.). REI. https://www.rei.com/learn/expert-advice/batteries.html
[29] What is Google's Gemini AI tool (formerly Bard)? Everything you need to know. (n.d.). ZDNET. https://www.zdnet.com/article/what-is-googles-gemini-ai-tool-formerly-bard-everything-you-need-to-know/
[30] Battery comparison chart. (n.d.). Energizer. https://energizer.com/batteries/battery-comparison-chart/
[31] *Duracell Optimum - Duracell Batteries | AA, AAA, Rechargeable, Coin Button*. (n.d.). Duracell. https://www.duracell.com/en-us/product/duracell-optimum-battery/
[32] *Energizer MAX® Alkaline Batteries*. (n.d.). Energizer. Retrieved April 10, 2024, from https://energizer.com/energizer-max-alkaline-batteries/
[33] *AA HIGH ENERGYTM Alkaline Batteries*. (n.d.). Rayovac. Retrieved April 10, 2024, from https://rayovac.com/product/aa-high-energy-alkaline-batteries/

[34] Ashtari, H. (2023, March 13). *Unix vs. Linux vs. windows: How they compare: Spiceworks - Spiceworks*. Spiceworks Inc. https://www.spiceworks.com/tech/tech-101/articles/unix-linux-windows-comparison/
[35] G., D. (2023, December 1). *Centos vs ubuntu - which one to choose for your web server*. Hostinger Tutorials. https://www.hostinger.com/tutorials/centos-vs-ubuntu
[36] Dillon, L. (2020, July 15). *CentOS vs. ubuntu: Feature comparison and use cases*. OpenLogic by Perforce. https://www.openlogic.com/blog/centos-vs-ubuntu
[37] Bonuccelli, G. (2022, February 15). *Cloud vs server: Learn the key differences and benefits*. Server and Cloud Blog. https://www.parallels.com/blogs/ras/cloud-vs-server/
[38] Wassim. (2023, September 13). *Comparing desktop application development frameworks: Electron, flutter, Tauri, react native, and...* Medium. https://medium.com/@maxel333/comparing-desktop-application-development-frameworks-electron-flutter-tauri-react-native-and-fd2712765377
[39] How HDMI Works. (2007, October 8). HowStuffWorks. https://electronics.howstuffworks.com/hdmi.htm#:~:text=HDMI%20uses%20transition%20minimized%20differential%20signaling%20%28TMDS%29%20to
[40] Ylonen, T. (1996). SSH - Secure login connections over the Internet. *Proceedings of the 6th USENIX Security Symposium*, 37-42. USENIX.
[41] *ANSI and IEC battery standardization nomenclature*. (n.d.). Www.microbattery.com. https://www.microbattery.com/battery-standardization-nomenclature
[42] *Projector Specs Explained – Your Ultimate Projector Buyer's Guide! (2024)*. (2023, September 12). https://theaterdiy.com/projector-specs-explained-ultimate-projector-buyers-guide/
[43] *Camera Measurement Standards*. (n.d.). Www.imaging.org. Retrieved April 10, 2024, from https://www.imaging.org/IST/IST/Standards/Camera_Measurement_Standards.aspx
[44] Roller, J. (2023, December 27). Ethical Pros and Cons of AI Image Generation. IEEE Computer Society. https://www.computer.org/publications/tech-news/community-voices/ethics-of-ai-image-generation
[45] Ortiz, S. (2023, September 15). What is ChatGPT and why does it matter? Here's everything you need to know. ZDNET; ZDNET. https://www.zdnet.com/article/what-is-chatgpt-and-why-does-it-matter-heres-everything-you-need-to-know/
[46] What is Copilot (formerly Bing Chat)? Here's everything you need to know. (n.d.). ZDNET.

https://www.zdnet.com/article/what-is-copilot-formerly-bing-chat-heres-everything-you-need-to-know/
[47] *Raspberry Pi Documentation - Raspberry Pi Hardware*. (n.d.). Www.raspberrypi.com. https://www.raspberrypi.com/documentation/computers/raspberry-pi.html
[48]LSM0805 Series 0805 SMD LED Package 2.0 x1.25x 0.8 mm SMD Chip LED LSM0805412V Red SMD LED. Low Profile Surface Mount LED with High intensity light output and low power consumption Application Key Features. (n.d.). Retrieved April 19, 2024, from https://vcclite.com/wp-content/uploads/2019/05/LSM0805412V-Red-Datasheet.pdf
[49] Raspberry Pi Documentation - Raspberry Pi Hardware. (n.d.). Www.raspberrypi.com. https://www.raspberrypi.com/documentation/computers/raspberry-pi.html
[50]"OpenCV: Camera Calibration with OpenCV." *Docs.opencv.org*, docs.opencv.org/3.4/d4/d94/tutorial_camera_calibration.html.
[51] How to Draw Pusheen with Donut 🍩 Easy Drawing on a Whiteboard. (2019, November 8). www.youtube.com. https://www.youtube.com/watch?v=zxxJWL2TZEA
[52] Datta, S., Ku, A., Ramachandran, D., & Anderson, P. (2023, December 27). *Prompt expansion for adaptive text-to-image generation*. arXiv.org. https://arxiv.org/abs/2312.16720
[53] lllyasviel. (2023, May 16). ControlNet 1.1. GitHub.
[54] *Canny edge detection*. OpenCV. (n.d.). https://docs.opencv.org/4.5.2/da/d22/tutorial_py_canny.html
[55] *CV::Linesegmentdetector class reference*. OpenCV. (n.d.-b). https://docs.opencv.org/4.5.2/db/d73/classcv_1_1LineSegmentDetector.html
[56] Xie, S., & Tu, Z. (2015, October 4). *Holistically-nested edge detection*. arXiv.org. https://arxiv.org/abs/1504.06375
[57] Most Efficient PCB Solutions for engineers and hobbyists - JLCPCB. (n.d.). Jlcpcb.com. https://jlcpcb.com/aboutUs
[58] *DATASHEET Raspberry Pi 4 Model B Release 1*. (2019). https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-datasheet.pdf
[59] jira about. (n.d.). Bing. Retrieved March 10, 2024, from https://www.bing.com/search?q=jira+about&qs=n&form=QBRE&sp=-1&pq=jira+about&sc=11-10&sk=&cvid=5CB8D38767B8473DB9907FE67C7F3538&ghsh=0&ghacc=0&ghpl=

Appendix C - Copyright Permissions

Request For Reference Permission

Francisco Soriano
To: asandvik@calpoly.edu

Reply | Reply all | Forward

Thu 1/18/2024 3:48 PM

Good afternoon,

Hello, my name is Francisco Soriano and I am taking a Senior Design capstone course as a Senior Computer Engineering student at the University of Central Florida.

I would like to respectfully request permission to use your work - "Dynamic Projection Mapping" in our "Related Works" section in my Senior Design Project documentation report. Is this okay with you?

Best,
Francisco Soriano

Francisco Soriano
AMD Research Fellow
University of Central Florida - Computer Engineering
321-603-9123

Figure C.1 *Copyright Request Email Screenshot*

Retention: UCF Delete after 10 Years (10 years) Expires: Sun 1/15/2034 9:57 PM

AS

Addison Sandvik <asandvik@calpoly.edu>
To: Francisco Soriano

Thu 1/18/2024 9:57 PM

Hi Francisco,

Yes, you may use my work as a reference.

Good luck!
Addison Sandvik

...

Figure C.2 *Copyright Response Email Screenshot*

VM

Victoria Moreno <vicmoreno21@outlook.com> Today at 8:10 AM

To: lvmin@cs.stanford.edu

Hello,

I hope you are doing well. My name is Victoria Moreno and I am taking a Senior Design capstone course as a Senior Computer Engineering student at the University of Central Florida.

I would like to respectfully request permission to use your work - "Adding Conditional Control to Text-to-Image Diffusion Models" in our "Related Works" in my Senior Design Project documentation report. Is this okay with you?

Best,
Victoria Moreno

Figure C.3 *Copyright Request Email Screenshot*

Liana Ez drawing
liana_e_z_drawing · Instagram
View profile

Thu 5:33 AM

Good morning,

Hello, my name is Declan, and I am taking a Senior Design capstone course as a Senior Computer Engineering student at the University of Central Florida. I'm working on a generative projection mapping project. As part of our work, I came across your image of "Draw Pusheen with Donut," loved it, and I am looking to include it as an example image in our work to demonstrate our AI generation tool on simple sketches.

I would like to respectfully request permission to use your work regarding your "Draw Pusheen with Donut" video/created image in our "Related Works" section in my Senior Design Project documentation report. Is this okay with you?

Best,
Declan Carter

Additionally, would we be able to record this conversation over email for my course? My email is declan@ucf.edu. Thank you!

https://www.youtube.com/watch?v=zxxJWL2TZEA (from this video for reference)

Figure C.4 *Copyright Request Screenshot*
(Email Could Not Be Located)

Appendix D - Code

```
# Declan Carter; UCF '24
# Senior Design 1; GPMS Prototype
# Description:
#       This script configures and uses Picamera2 for capturing video
streams and applies real-time edge detection.
#       It features a dual-window setup for point selection and edge
detection visualization.
#       Users can interactively select corner points on a white screen
which will then be used to warp the perspective
#       of the captured frames. The script also allows for dynamic
adjustment of edge detection thresholds via keyboard inputs.
# Usage:
#       1. Run the script in an environment that supports cv2 and
Picamera2.
#       2. Drag the 'Point Selection' window to the desired screen as
instructed by the console.
#       3. Select points on the 'Point Selection' window by clicking the
left mouse button.
#       4. Press 'u' to capture the current frame and use it for edge
detection.
#       5. Adjust the hi and lo edge detection thresholds using 'o'/'p'
and 'k'/'l' keys respectively.
#       6. Press 'q' to quit the script.
#       7. Additional functions include resetting selected points ('r'),
applying a watermark ('c'), and resetting thresholds to default ('d').

import cv2
from picamera2 import Picamera2, Preview
import numpy as np
import datetime
import time

# Initialize Picamera2
picam2 = Picamera2()

# Capture the Resolution
width = picam2.sensor_resolution[0]
height = int(width * (9/16))  # The display will be 16:9

# Set Image Resolution
picam2.preview_configuration.main.size = (width, height)
picam2.preview_configuration.main.format = "RGB888"
picam2.preview_configuration.align()
picam2.configure("preview")
picam2.start()

# Create Window 2
cv2.namedWindow('Edge Detection', cv2.WINDOW_NORMAL)
cv2.setWindowProperty('Edge Detection', cv2.WND_PROP_FULLSCREEN,
cv2.WINDOW_FULLSCREEN)

# Create Window 1, wait for user to drag it to display 2
white_screen = np.full((600, 800, 3), 255, dtype=np.uint8)
cv2.namedWindow('Point Selection', cv2.WINDOW_NORMAL)
```

```python
cv2.imshow('Point Selection', white_screen)
# Display instruction
print("Please drag the window to the desired screen and press any key to
continue...")

cv2.waitKey(0)  # Wait for user input

# Default edge detection thresholds
low_threshold = 100
high_threshold = 200
default_low_threshold = low_threshold
default_high_threshold = high_threshold

# Default corner points and state
default_pts_webcam = np.float32([[0, 0], [0, 720], [1280, 720], [1280,
0]])
new_pts = []
reset_mode = True
matrix = None

# Mouse callback function
def select_point(event, x, y, flags, param):
    global new_pts
    if event == cv2.EVENT_LBUTTONDOWN and len(new_pts) < 4:
        new_pts.append([x, y])

# Set mouse callback
cv2.setMouseCallback('Point Selection', select_point)

# Function to reset the corner points
def reset_points():
    global new_pts, reset_mode
    reset_mode = True
    new_pts = []
    cv2.setMouseCallback('Point Selection', select_point)

# Function to calculate the perspective transformation matrix
def get_matrix():
    pts_projection = np.float32([[0, 0], [0, 720], [1280, 720], [1280,
0]])
    return cv2.getPerspectiveTransform(np.float32(new_pts),
pts_projection)

# Function to apply Canny edge detection
def apply_edge_detection(frame, low, high):
    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
    blurred = cv2.GaussianBlur(gray, (5, 5), 0)
    edges = cv2.Canny(blurred, low, high)
    return edges

# Function to add watermark to the frame
def add_watermark(frame, text):
    font = cv2.FONT_HERSHEY_SIMPLEX
    cv2.putText(frame, text, (10, 700), font, 1, (255, 255, 255), 2,
cv2.LINE_AA)
    return frame
```

```
# Function to capture and save frame with timestamp
def capture_frame_with_timestamp(frame):
     print("Attempting to save the frame:", type(frame), frame.shape if
isinstance(frame, np.ndarray) else "Not a frame")
   if isinstance(frame, np.ndarray) and frame.size != 0:
       timestamp = datetime.datetime.now().strftime("%Y%m%d_%H%M%S")
       filename = f'captured_frame_{timestamp}.jpg'
       cv2.imwrite(filename, frame)
       print(f"Frame captured and saved as {filename}.")
   else:
       print("Error: No valid frame to save.")

def capture_frame():
   frame = picam2.capture_array()
   if frame is None:
       print("Failed to capture video.")
       return None

   # Flip the frame for both point selection and edge detection
   frame = cv2.flip(frame, 0)
   frame = cv2.flip(frame, 1)

   return frame



while True:
   frame_ps = capture_frame()

   if reset_mode:
       if new_pts != []:
           for pt in new_pts:
               cv2.circle(frame_ps, tuple(pt), 5, (0, 255, 0), -1)
       if len(new_pts) == 4:
           matrix = get_matrix()
            still_frame = frame_ps.copy()  # Capture still frame for edge
detection
           reset_mode = False
       frame_ed = np.full((600, 800, 3), 255, dtype=np.uint8)
   else:
       if matrix is not None and still_frame is not None:
            transformed = cv2.warpPerspective(still_frame, matrix, (1280,
720))
              frame_ed = apply_edge_detection(transformed, low_threshold,
high_threshold)
       else:
           frame_ed = np.full((600, 800, 3), 255, dtype=np.uint8)

   cv2.imshow('Point Selection', frame_ps)
   cv2.imshow('Edge Detection', frame_ed)

   key = cv2.waitKey(1) & 0xFF
   if key == ord('q'):
       break
   elif key == ord('o'):
       high_threshold = max(high_threshold - 10, 0)
   elif key == ord('p'):
```

```
        high_threshold = min(high_threshold + 10, 255)
    elif key == ord('k'):
        low_threshold = max(low_threshold - 10, 0)
    elif key == ord('l'):
        low_threshold = min(low_threshold + 10, 255)
    elif key == ord('r'):
        reset_points()  # Reset the points
          frame_ed = np.full((600, 800, 3), 255, dtype=np.uint8)   # Edge
detection shows white
    elif key == ord('u'): # Update Frame being used for edge detection
        # Flash White Frame to Capture Image
        frame_ed = np.full((600, 800, 3), 255, dtype=np.uint8)
        cv2.imshow('Edge Detection', frame_ed)
        cv2.waitKey(3000)  # Wait for 3 seconds with GUI responsive
        still_frame = capture_frame()
    elif key == ord('c'):
        watermark_text = f'Low: {low_threshold}, High: {high_threshold}'
                  frame_with_watermark   =   add_watermark(frame_ps.copy(),
watermark_text)
        capture_frame_with_timestamp(frame_with_watermark)
        print("Frame captured with watermark.")
    elif key == ord('d'):
        low_threshold = default_low_threshold
        high_threshold = default_high_threshold
        print("Thresholds reset to default values.")

# Release the webcam and destroy all OpenCV windows
picam2.stop()
cv2.destroyAllWindows()
```

Appendix E - Figures

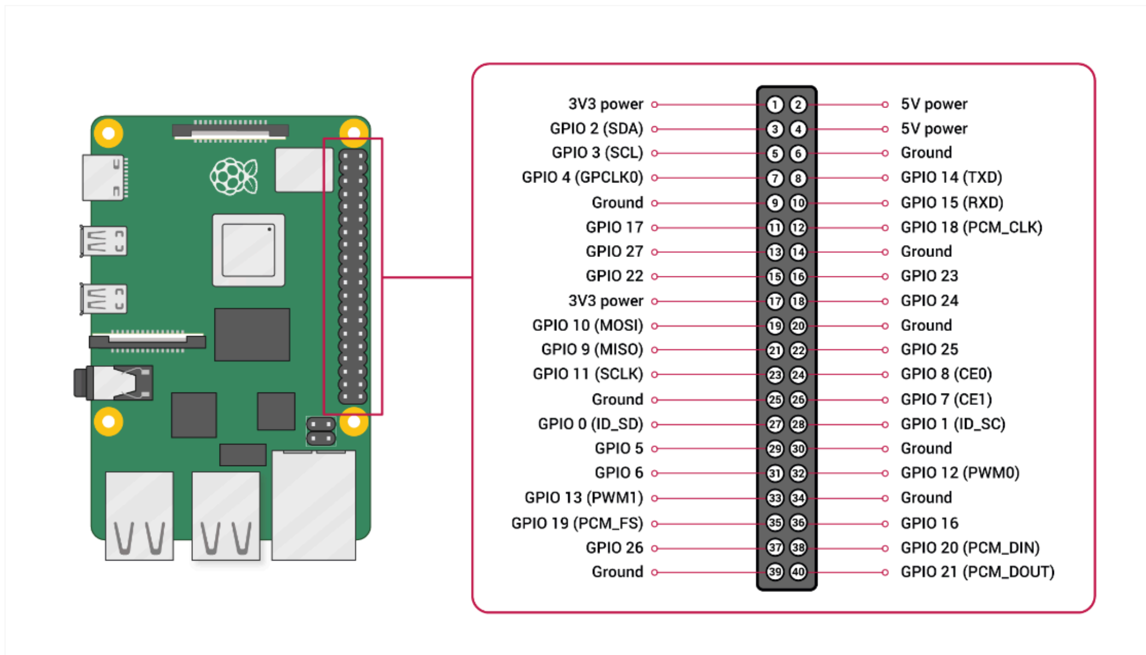


Figure E.1 *Raspberry Pi 4B GPIO Layout* [47]

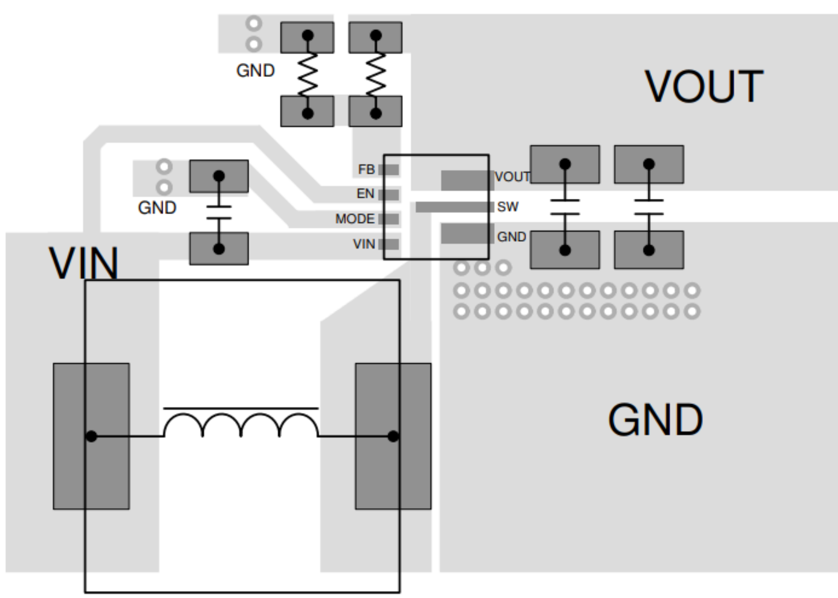


Figure E.2 *Recommended Regulator Layout* [26]

Appendix F - Tables

The BOM table below is specifically for the custom PCB. All of the components were ordered on DigiKey Electronics.

Part	Value	Description	Price (USD)
C1	10μF	CAPACITOR, American symbol	0.1
C2	47μF	CAPACITOR, American symbol	0.26
C3	47μF	CAPACITOR, American symbol	0.26
D1	LED_RED	N/A	0.29
L1	XAL4020-102ME BIND_XAL4020_ 4X4_COC-L	N/A	2.55
R1	100kΩ	RESISTOR, American symbol	0.1
R2	732kΩ	RESISTOR, American symbol	0.1
R3	27Ω	RESISTOR, American symbol	0.1
SV1	N/A	PIN HEADER	0.13
SW1	1825255-1SW3_1 825255-1_TEC	N/A	1.18
U$1	BATTERY-HOLD ER-2AA	Battery Holder 2 x AA	1.11
U$2	BATTERY-HOLD ER-2AA	Battery Holder 2 x AA	1.11
U2	TPS61022RWUR RWU0007A	N/A	2.15
TOTAL BEFORE	**N/A**	**N/A**	**9.44**

SHIPPING			
TOTAL AFTER SHIPPING	**N/A**	**N/A**	**34.84**

Table F.1 *Custom PCB BOM*

The BOM table below is specifically for the AMD machine. AMD generously lended the authors compute time on the AMD machine.

ITEM	QUANTITY	PRICE ESTIMATE (USD)
CPU*	1	516.54
CPU Cooler*	1	128.64
GPU*	1	1,049.99
Motherboard*	1	259.99
RAM*	1	349.99
Power Supply*	1	489.32
Case*	1	139.99
Storage OS*	1	159.99
Storage APPS*	1	149.99

Table F.2 *AMD Machine BOM*