Design Review Report

ABSTRACT:

EmotiSense is a real-time emotion detection product designed to assist autistic and neurodivergent individuals in recognizing and interpreting their conversation partner's emotions. Our product uses a webcam to capture live video, which is then run through our in-house machine-learning model to detect emotions based on captured facial expressions. The detected emotions are displayed on an iPad in real-time and conveyed through haptic and visual feedback via a wearable bracelet. EmotiSense aims to enhance interpersonal communication and emotional understanding for individuals with autism, improving their social interactions and overall well-being.

INTRODUCTION:

Emotional awareness and recognition are key elements to successful communication. For individuals with autism or other neurodivergent conditions, recognizing the emotions of others can be a significant challenge. This undoubtedly harms their ability to build interpersonal relationships and engage in effective communication. Worse, these communication problems can lead to frustration for both the neurodivergent individual and their partner who often is misunderstood. EmotiSense is designed to address these issues by providing real-time emotion recognition feedback during conversations, enabling users to better understand the emotional cues of their conversation partners.

EmotiSense combines computer vision and machine learning to detect emotions from facial expressions captured via a webcam live video stream. It processes this data on a Nvidia Jetson which can provide real-time feedback through both visual displays on an iPad and haptic feedback on a wearable bracelet. The goal of EmotiSense is to bridge the emotional recognition gap for individuals with autism giving them a means to create connections like neurotypical individuals. This fosters more meaningful and comfortable social interactions for all involved. By offering a non-intrusive, real-time solution, EmotiSense promotes improved communication, better social understanding, and enhanced well-being for its users.

USE-CASE REQUIREMENTS:

The EmotiSense bracelet is designed to provide neurodivergent individuals with real-time feedback to aid in interpreting emotional cues. To ensure the system is both functional and user-friendly, the following requirements have been established:

The first is clear and intuitive feedback with easily distinguishable feedback. User studies should demonstrate that 90% of participants can correctly identify the type of feedback (either haptic or visual) associated with specific emotions within 3 seconds of actuation. This metric ensures that the bracelet's signals are intuitive enough to help users quickly recognize changes in emotional states, thus enhancing their understanding of social interactions.

Another use case requirement for the user is non-intrusive wearability. We want the device to be comfortable from the user's perspective and wearable for an extended period of time. To validate the bracelet's wearability, it should be rated at least 7 out of 10 for comfort by users during testing sessions. In addition, the device should weigh less than 250 grams and smaller than 3" x 3" x 3" to ensure discrete conversations and to allow the device to be used in various social situations without drawing unwanted attention.

Additionally, we want accurate emotional cues with higher confidence. This means that feedback from the bracelet should only be provided when the emotion detection algorithm achieves a confidence level of at least 80%. This requirement minimizes the likelihood of false positives, which could mislead the user or cause anxiety. During testing, the system must demonstrate an error rate of less than 10% for incorrectly signaling emotions, thus maintaining a high standard of accuracy that is crucial for reliable emotional assistance.

Lastly, the bracelet should support at least three levels of vibration intensity and four distinct LED colors corresponding to different emotions, enabling users to customize the feedback to their preferences. Usability assessments should show that 80% of participants find the customization options satisfactory or above, indicating that the system provides a suitable degree of adaptability for different user needs.

In terms of the machine learning and computer vision components, we have concluded that the emotion detection model should achieve an accuracy of 70-75%. Given that the best emotion recognition models operate with an accuracy in the range of 75-80% (Beltramin, 2023), we believe that we should be able to reach an accuracy slightly below that range. An accuracy of 65-75% matched the best emotion recognition models around 5 years ago which should be obtainable with our in-house model given the additional tools and computing power of the Nvidia Jetson today.

The second use-case requirement related to the machine learning model focuses on the feedback time. The system must deliver feedback within 300 milliseconds from the moment an emotion is detected. According to the ITU-T G.114 standard, response times under 150 milliseconds are typically imperceptible to users, ensuring natural interaction flow. However, latencies up to 300 milliseconds are still considered acceptable in many real-time systems, including human-machine interactions, without significantly disrupting conversational dynamics. Research shows that delays beyond 300ms can lead to noticeable disruptions, causing feedback to feel delayed and less effective for real-time support.

By keeping the feedback within this limit, the system ensures that users can adjust their behavior in near real-time, preserving the natural flow of conversation. Although 300ms is functional, aiming for 100 milliseconds would make the feedback imperceptible, providing a smoother and more seamless user experience. We will continue this discussion in the design requirements section to highlight what level of real-time feedback we are capable of providing.

Without any assistance, individuals with autism typically take over 4100 milliseconds to gauge emotions with a 73% accuracy rate (Drimalla et al., 2021) for most emotions with accuracy rates as low as 50% for emotions like sadness, anger, and fear (Bleier et al., 2024). Given these low and slow rates of emotion detection, targeting a model accuracy of 70-75% and a feedback time of 300 milliseconds is appropriate for our product. EmotiSense dramatically reduces the time needed to interpret emotions while maintaining similar accuracy. Additionally, with feedback arriving over 13 times faster than an autistic individual's average response time, EmotiSense offers real-time support during conversations, helping users react more quickly and consistently to emotional cues without overwhelming them. This balance ensures both timely and effective assistance in social interactions.

The system must also achieve a level of accuracy that helps users better understand their conversation partner’s emotions. The goal is for the feedback to significantly improve users' emotional recognition compared to their baseline abilities, particularly for individuals with autism, who often find it challenging to recognize emotions like anger and sadness. To ensure its usability, the system must detect a wide enough array of emotions. At a minimum, the system must recognize four key emotions—happy, sad, angry, and neutral—through the bracelet. For users seeking more comprehensive emotional insights, the iPad interface will offer feedback for six core emotions—happy, sad, angry, neutral, fear, and surprise—giving users the flexibility to choose between simplified or more detailed feedback based on their needs.

The EmotiSense bracelet is designed to detect four core emotions—happy, sad, angry, and neutral—based on the specific difficulties autistic individuals face with emotion recognition. As stated before, research shows that emotions like anger and sadness are particularly challenging for individuals with autism to identify accurately (Bleier et al., 2024); however, they are some of the most important emotions to recognize during a conversation. By focusing the bracelet on these core emotions, we ensure that the feedback is clear, high-confidence, and easy to interpret, helping users quickly respond to crucial emotional cues without being overwhelmed.

On the iPad, the system expands to detect six core emotions, including fear and surprise, providing additional context for users who want more detailed insights. This broader set allows for more nuanced feedback without adding complexity to the real-time interaction. The combination of four emotions on the bracelet and six on the iPad strikes a balance, offering enough emotional range to be useful while keeping the system user-friendly and effective in assisting autistic individuals in recognizing emotions during social interactions.

The iPad display requirements focus on user experience. We want to convey as much information to the user as possible, without becoming a distraction that will interrupt the flow of conversation. Some additional data that is included in the web app display is classification confidence (as a percentage), and emotion duration in seconds (how long has the user been displaying this emotion as interpreted by the model). Users must be able to have the device in their field of vision, glance at it occasionally without detracting from the conversation. For this reason we will asks users to rate the web app experience, and give rating/feedback on the display itself. Users in user testing will evaluate distraction, communication, and usability, while conversation participants will evaluate the disruption that the device caused from their point of view.

ARCHITECTURE AND/OR PRINCIPLE OF OPERATION

One sub architecture of the EmotiSense tech stack is the embedded bracelet device. The purpose of the wearable device is to provide haptic and visual feedback to users. The feedback helps users interpret emotional cues from their conversational partners. With this in mind, the design emphasizes relatability and efficient operation while maintaining ease of use and discretion.

The haptic feedback mechanism is powered by a Mini Vibrating Motor Disc and Haptic Motor Controller managing the vibration control. This setup allows for smooth, gradual vibration feedback, rather than binary on/off signals, giving the user a more intuitive understanding of emotional intensity. The motor provides tactile sensations when a significant emotional state change is detected. By modulating the intensity of the vibrations, the system can communicate emotional changes more clearly, helping neurodivergent individuals understand not only what emotion is present but also how confident the model's prediction is.

The visual feedback system consists of a NeoPixel Jewel, which contains individually addressable RGB LEDs. These LEDs change color depending on the emotion detected by the machine learning model running on the Nvidia Jetson. The bracelet is designed to light up only when the system has high confidence in the detected emotion, preventing erroneous signals that could confuse the user. This high-confidence feedback ensures that the emotional cues are accurate and that the user can rely on the bracelet's signals without hesitation.

The bracelet receives data from the Nvidia Jetson, which processes the incoming video frames captured by the connected webcam. The Jetson, running a custom convolutional neural network (CNN), identifies the emotion of the conversational partner in real-time. Currently, the Jetson communicates with the Adafruit Feather via a UART (Universal Asynchronous Receiver-Transmitter) connection, providing fast, low-latency data transfer between the devices. This ensures that the bracelet actuates within 150 milliseconds after the emotion is detected, maintaining synchronization with the conversation's emotional flow. However, this UART communication is part of the initial implementation, and the eventual goal is to transition to Bluetooth Low Energy (BLE) for improved wireless connectivity and greater flexibility in real-world usage scenarios. BLE will allow for a more streamlined and user-friendly experience without the need for wired connections.

In addition to ensuring low-latency communication and precise feedback, the bracelet is designed with comfort and usability in mind. It is lightweight and adjustable, making it suitable for a range of users. The haptic feedback is subtle yet clear, ensuring that it can be felt without being distracting or uncomfortable, while the visual feedback is designed to be noticeable without being overwhelming. Ethical considerations in the design, such as minimizing false positives in emotional detection, are critical to ensure that the device provides meaningful assistance to users without causing frustration.

The computer vision and emotion recognition model is being built using existing facial recognition technologies. Currently, I have settled on the Haar Cascade classifier which is a commonly used model for facial detection. Using the Haar Cascade model, our system detects the conversational partner's face in real-time video frames, isolates it, and then crops the image to focus specifically on the facial region. This cropped facial image is then passed through a series of custom preprocessing steps before being piped in to train our custom emotion recognition model. These preprocessing steps include scaling the images to a standard size, normalizing pixel values to improve model performance, and applying data augmentation techniques such as rotation, flipping, or contrast adjustments. These methods ensure the model is robust against variations in lighting, orientation, facial expressions, race, and many other confounding factors. This provides consistent input for accurate and unbiased training. This preprocessing pipeline helps enhance the model's ability to generalize across diverse facial data, resulting in more reliable emotion classification in real-world conditions.

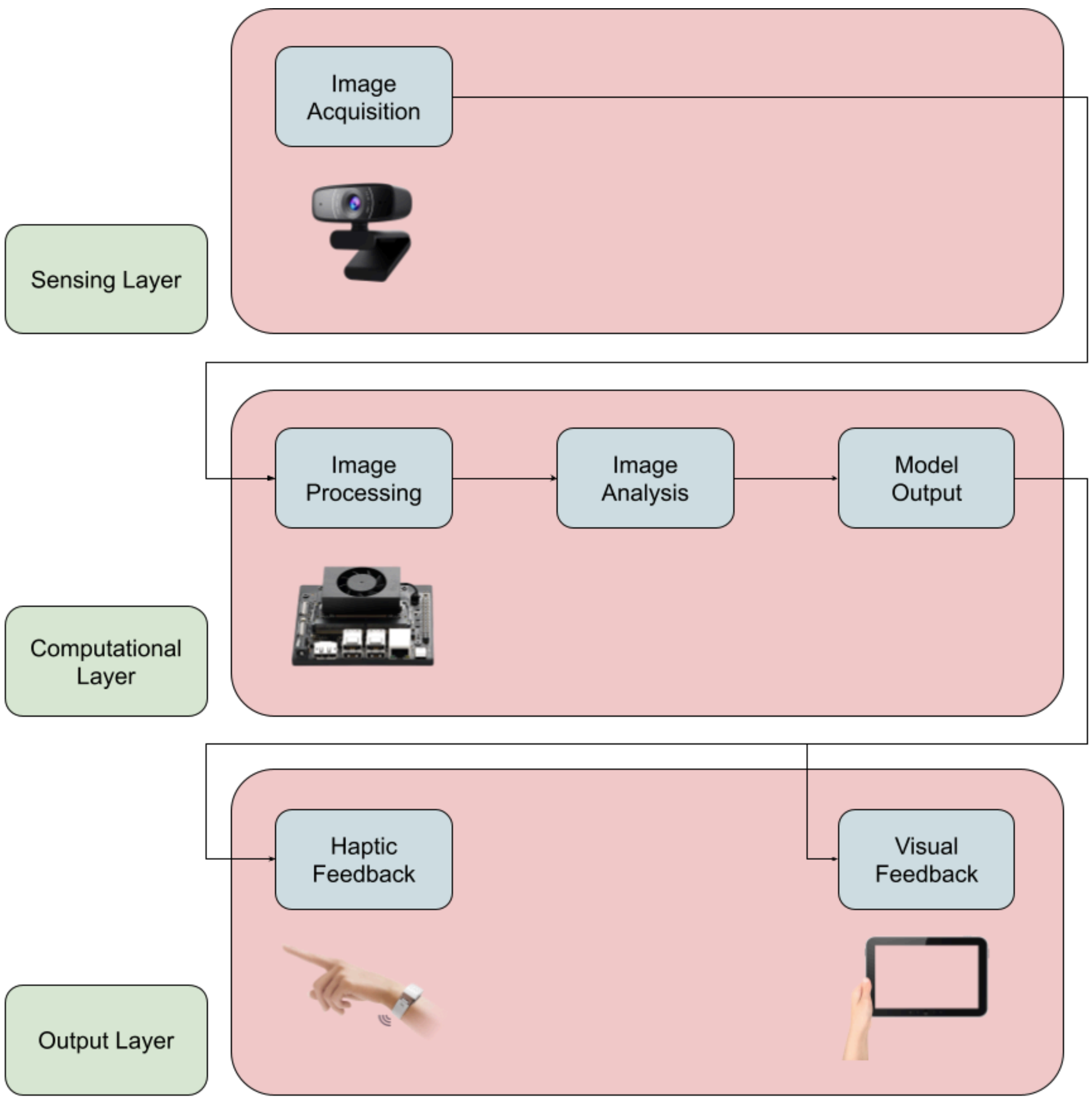
OpenCV is utilized for integrating this facial detection capability. Once faces are detected, a custom-trained emotion recognition CNN is built on top of this to classify emotional expressions. The emotion recognition CNN is trained on datasets like FER-2013, which provides labeled data for six core emotions, allowing the system to accurately classify emotional states in real-time.

I plan to develop and train this emotion recognition model using Google Colab, which provides access to GPUs for quicker training. The Colab environment allows for experimenting with hyperparameters, data augmentation techniques, and optimizing the CNN architecture without the limitations of local hardware. Once trained, the model is optimized for deployment on the Nvidia Jetson using TensorRT and CUDA (Nvidia's parallel computing platform), ensuring that the model runs efficiently and with minimal latency. This approach allows the system to maintain real-time performance while leveraging the computational power of the Jetson for local inference during operation.

To ensure confidence in the emotion detection before providing feedback to the user, the outputted emotions will be averaged over a short time window. This averaging smooths out potential misclassifications or momentary shifts in facial expressions, preventing false positives from triggering feedback unnecessarily and distracting our users. By aggregating the detected emotions over several frames, the system establishes a more stable emotional state, and only when a consistent emotional trend is identified will the system trigger the bracelet's haptic or visual feedback. This approach ensures that the user is alerted to emotionally significant cues only when the model is confident, avoiding interruptions caused by fleeting or ambiguous facial expressions.

Once finalized, the model will operate by capturing real-time video through the Logitech webcam in conjunction with OpenCV, then feeding the detected faces into the emotion recognition model. The model processes these inputs in real time, classifying the detected emotions, which are then transmitted to the bracelet and iPad for immediate feedback. The modular nature of this architecture allows for future improvements such as averaging our custom emotion recognition model with heavier-weigh industry models.

ARCHITECTURE DIAGRAM:



DESIGN REQUIREMENTS:

EmotiSense is designed to provide real-time feedback to users with neurodivergent conditions, particularly individuals with autism, who may struggle with recognizing emotions during conversations. The primary requirement for the system is minimizing the delay between when a conversational partner's emotion is detected and when that information is relayed to the user. For the bracelet, feedback must be provided within 150 milliseconds. This delay is broken down into 20 ms for the camera to capture the frame, 15 ms for data transmission via USB, 50 ms for the Jetson to process the image and run the emotion detection model, 50 ms for the bracelet to actuate through its haptic and visual components, and an additional 15 ms as buffer time . This fast response time is essential to maintain the real-time nature of the feedback, allowing users to adjust their behavior quickly based on their conversational partner's emotional responses. Similar latency requirements have been explored in real-time emotion recognition systems, where low-latency feedback is crucial for maintaining natural conversation flow without interruptions (Barros et al., 2015).

For the iPad display, the feedback time is extended to 400 milliseconds to account for an additional 300 ms of WiFi transmission time between the Jetson and the iPad. This slightly longer delay is still within acceptable limits for real-time interaction, ensuring that the visual feedback remains timely and relevant to the conversation. Prior research on wearable emotion recognition systems supports the idea that delays of up to 500 ms are acceptable for providing real-time feedback without disrupting communication dynamics (Zhou et al., 2019). The EmotiSense system's latency goals are aligned with these findings, ensuring that both the bracelet and iPad provide feedback to the user without perceptible delay.

The accuracy of the machine learning model used for emotion detection is another critical consideration. The system is designed to achieve an accuracy of 70-75%, which is competitive with the current state of the art in facial emotion recognition models. The best models, such as those based on convolutional neural networks (CNNs), typically achieve accuracies between 75-80% under ideal conditions (Li et al., 2018). However, research shows that autistic individuals often perform worse at recognizing emotions like anger and sadness, which are particularly challenging (Rutherford et al., 2002). By setting the system's accuracy goal slightly lower than the best-performing models, EmotiSense accounts for the real-world variability in input data and the difficulty of reliably detecting emotions in dynamic environments. Moreover, this accuracy is sufficient to provide meaningful feedback, as human emotion recognition in natural settings is not perfect, with typical accuracies ranging from 85-90% (Ekman, 1992).

Additionally, the system must operate at a minimum resolution of 28x28 pixels and maintain at least 30 frames per second (FPS) to ensure smooth real-time performance. Studies show that 30 FPS is standard in facial emotion recognition systems, as it allows for real-time tracking of facial expressions during dynamic interactions without lag, ensuring that subtle emotional cues are captured effectively. This frame rate is also common in FER datasets, such as CK+, which collects video sequences at 30 FPS for emotion recognition tasks (Lee et al., 2021). The resolution minimum of 28x28 pixels is chosen based on the FER-2013 emotion classification dataset that we are using which contains facial images with the same resolution. However, this is not a concern as the Logitech camera we are using will film at a resolution of 1920 x 1080 pixels. This will allow us to zoom in and center of the person's face even when not fully centered.

The bracelet feedback is deliberately designed to focus on only four emotion categories—happy, sad, angry, and neutral. This limitation is based on the need to minimize the risk of providing incorrect feedback, which could have negative effects on the user's confidence and social interactions. The bracelet only actuates when the system has high confidence in its prediction, avoiding low-confidence feedback that might confuse or mislead the user. In contrast, the iPad provides more comprehensive feedback by displaying six emotion categories—happy, sad, angry, neutral, fear, and surprise—along with a confidence score and emotion duration for each detected emotion. This approach allows users to see lower-confidence predictions while relying on the bracelet for more reliable, high-confidence feedback. The decision to limit the bracelet's feedback and provide more nuanced information on the iPad is consistent with the literature on assistive technology design, which emphasizes the importance of error minimization in systems for neurodivergent users (Dautenhahn et al., 2009). High confidence, clear feedback is crucial for these users to improve their social interactions without the risk of misinterpretation.

Finally, the system's data transmission must maintain low latency—below 300 milliseconds—to ensure smooth, real-time interaction between the Jetson, iPad, and bracelet. This latency ensures that the user experiences minimal delays between emotion detection and feedback, preserving the natural flow of conversation. Similar low-latency requirements have been noted in previous work on wearable assistive technologies, where delays of 200-300 ms are considered ideal for maintaining real-time responsiveness (Picard & Klein, 2002). Additionally, the system's battery life must provide a minimum of four hours of continuous operation, which is sufficient for most conversation settings and ensures usability in longer sessions, such as social gatherings or extended meetings. This requirement is typical for wearable assistive devices and allows for practical use over sustained periods without frequent recharging (Pentland, 2005).

DESIGN TRADE STUDIES

The design of the EmotiSense bracelet involves several trade-offs to meet the requirements for wearability, feedback accuracy, and real-time performance while maintaining user comfort and system reliability. One significant trade-off in the design is balancing power consumption with battery life. The 3.7V LiPo battery used in the bracelet has a limited capacity of 420mAh, which provides approximately 1.554Wh of energy. Lowering the brightness of the NeoPixel LEDs and reducing the intensity of the haptic feedback can extend battery life but may affect the clarity of the feedback, especially in bright environments or for users with reduced tactile sensitivity. Conversely, maximizing brightness and vibration intensity ensures clear feedback but may shorten the runtime to 2-3 hours. Thus, the design includes configurable settings that allow users to adjust feedback intensity according to their needs, offering a balance between usability and battery life.

The current implementation uses UART for communication between the Jetson and the Adafruit Feather, providing a simple and reliable wired connection. However, the goal is to transition to Bluetooth Low Energy (BLE) to enable a fully wireless user experience. This shift introduces trade-offs in terms of latency, power consumption, and complexity. While BLE allows for greater mobility and eliminates the need for a physical connection, it may introduce additional latency due to data transmission over the air. BLE's typical latency ranges from 10 to 50 ms, which could accumulate with other processing delays, potentially affecting the 150 ms responsiveness requirement. Additionally, BLE requires more sophisticated power management to avoid depleting the battery too quickly. To address these concerns, the transition to BLE will be tested against the UART-based baseline to ensure that latency remains within acceptable limits and that battery consumption does not exceed a 20% increase in average power usage.

Another design trade-off is the choice between haptic and visual feedback. The bracelet provides both types of feedback to cater to different user preferences and needs. However, there are inherent limitations and advantages to each type. Haptic feedback offers a discreet, non-visual way to communicate emotional changes and can be felt even when the user's attention is not on the bracelet. However, it may be challenging for users with reduced tactile sensitivity or in scenarios where vibration could be distracting (e.g., during meetings). On the other hand, visual feedback using the NeoPixel Jewel is more noticeable and can provide a clearer indication of emotion by displaying distinct colors. However, it requires the user to visually check the bracelet, which may not always be feasible in social interactions. The power consumption for LEDs is also higher than for the vibration motor, potentially reducing battery life. To optimize this trade-off, the system allows for the customization of feedback intensity and types. Users can select between different vibration levels and LED brightness, or disable one feedback type if not needed. This flexibility ensures that the bracelet can accommodate various user requirements without significantly compromising on battery life or clarity of feedback.

Finally, we have gone through many design changes and trade-offs during the creation of the computer vision and emotion recognition component of the project. Initially, we considered using existing emotion recognition models to quickly implement the system. While this approach would have provided a reliable baseline and saved development time, we decided against it in order to showcase the skills we have learned over our tenure at CMU. By building our own custom model, we could better demonstrate our proficiency in machine learning and tailor the model specifically to our project's needs. The tradeoff here was between speed of development versus the opportunity to apply advanced techniques and customize the model for our use case.

Next, we explored combining our custom model with an existing model and averaging their outputs to improve accuracy. While this could have increased robustness, we found that the added computational overhead and complexity in integrating two models would have made the system less efficient, harder to debug, and potentially lacking confidence in all predictions. The tradeoff here was accuracy versus system simplicity and potential performance. Ultimately, we decided that the benefits of combining models did not outweigh the additional complexity.

As we progressed, it became clear that creating everything from scratch would not be feasible within the project timeline. To stay on schedule, we opted to use the Haar Cascade facial recognition model from OpenCV for the face detection component. This decision allowed us to focus our efforts on the central task of emotion recognition while leveraging a proven, efficient model for face detection. The tradeoff was between fully custom development and ensuring we delivered a functional, real-time system within our deadline.

SYSTEM IMPLEMENTATION:

The EmotiSense bracelet is an integrated wearable system designed to provide real-time feedback on emotional states. The bracelet hardware comprises an Adafruit Feather microcontroller, a NeoPixel Jewel, a DRV2605L Haptic Motor Controller, an Adafruit Vibrating Mini Motor Disc, and a 3.7V 420mAh LiPo battery.

The microcontroller manages communication with the Jetson via UART, controls the feedback mechanisms, and regulates power distribution. The Adafruit Feather serves as the central control unit, handling the receipt of emotion data from the Jetson and controlling the haptic motor and LEDs based on that data. Operating at 3.3V, it is chosen for its low power draw, maximizing battery life while ensuring the system adheres to the 150 ms response time requirement.This estimate is based on the typical power consumption of the Adafruit Feather microcontroller, the DRV2605L Haptic Motor Controller, the Adafruit Vibrating Mini Motor Disc, and theNeoPixel Jewel. The 420mAh LiPo battery can supply approximately 1.554Wh of energy (3.7V * 420mAh). Given that the Feather draws approximately 9mA in idle mode and the NeoPixel LEDs and motor consume an additional 60mA to 150mA during operation, the total current draw would average around 120-180mA during regular use. This gives an estimated runtime of roughly 3-4 hours. This power management ensures that the system operates long enough for most social interactions, including casual conversations or extended meetings. The battery connects directly to the Adafruit Feather, which manages power distribution to all components, including the DRV2605L motor controller and NeoPixel Jewel, optimizing energy usage to meet the 4-hour requirement.

The DRV2605L Haptic Motor Controller drives the vibration motor, providing smooth, adjustable vibration patterns based on the confidence level of the emotion detected. This enables nuanced feedback that reflects the intensity of the emotional signal.

The NeoPixel Jewel, a compact LED array with seven individually addressable RGB LEDs, replaces the previous NeoPixel Ring. It displays distinct colors to represent different emotions, such as red for anger, blue for sadness, green for happiness, and yellow for neutral. The design change to the Jewel maintains compactness while providing sufficient visual feedback for emotion recognition.

The software running on the bracelet includes a real-time feedback control algorithm that processes incoming data from the Jetson and translates it into appropriate haptic and visual feedback. The software prioritizes consistency and clarity in feedback delivery to ensure that the bracelet's signals are easy to interpret and non-intrusive.

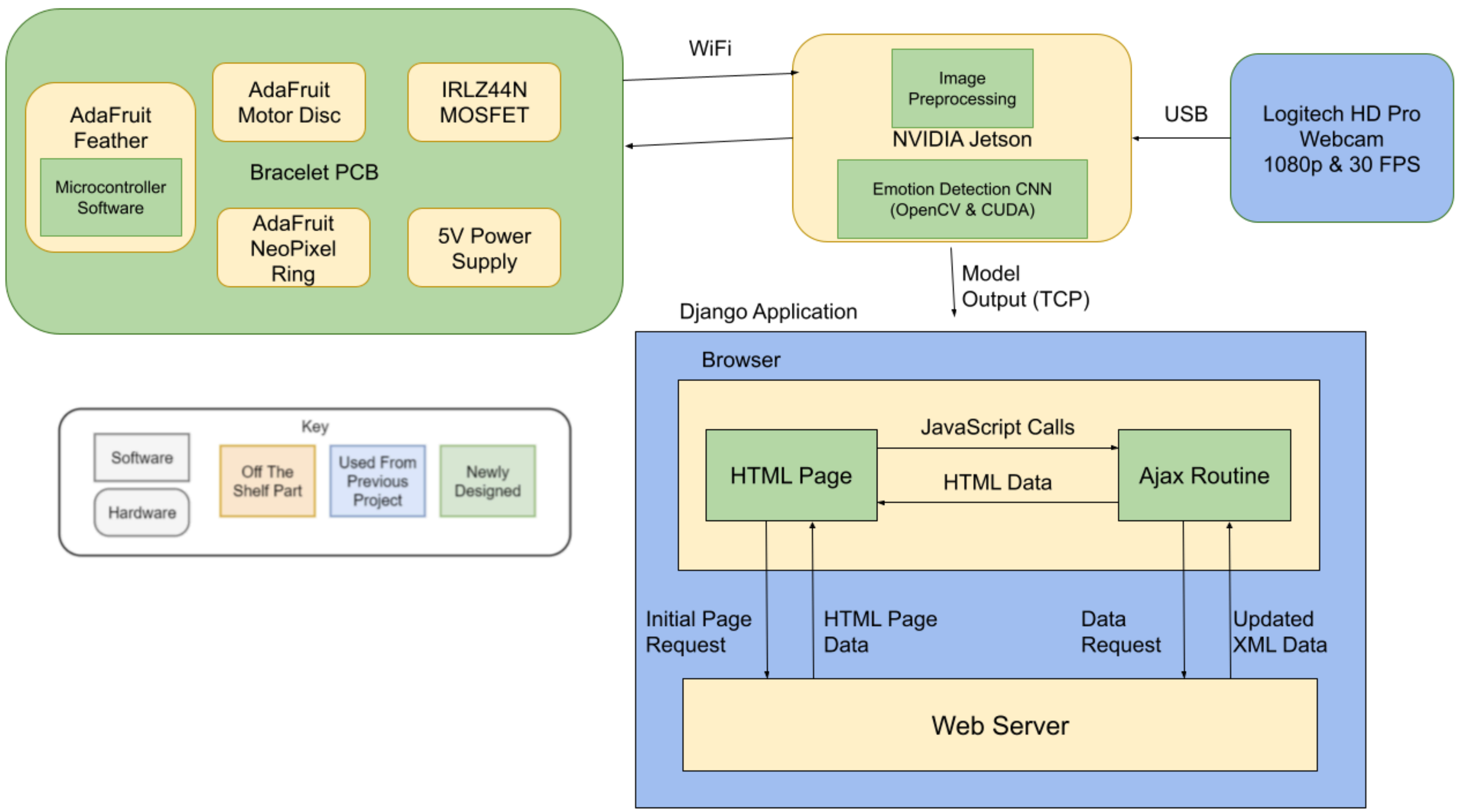
The Jetson sends emotion data to the Adafruit Feather via UART. Future iterations aim to replace UART with Bluetooth Low Energy (BLE) for wireless communication. A lightweight communication protocol is implemented to ensure minimal latency during data transmission. The software maps emotion confidence levels to specific haptic and visual feedback patterns. For high-confidence emotions, the vibration intensity and LED brightness are increased for clear feedback. For lower-confidence levels, feedback may be limited or omitted to avoid conveying uncertain signals to the user.

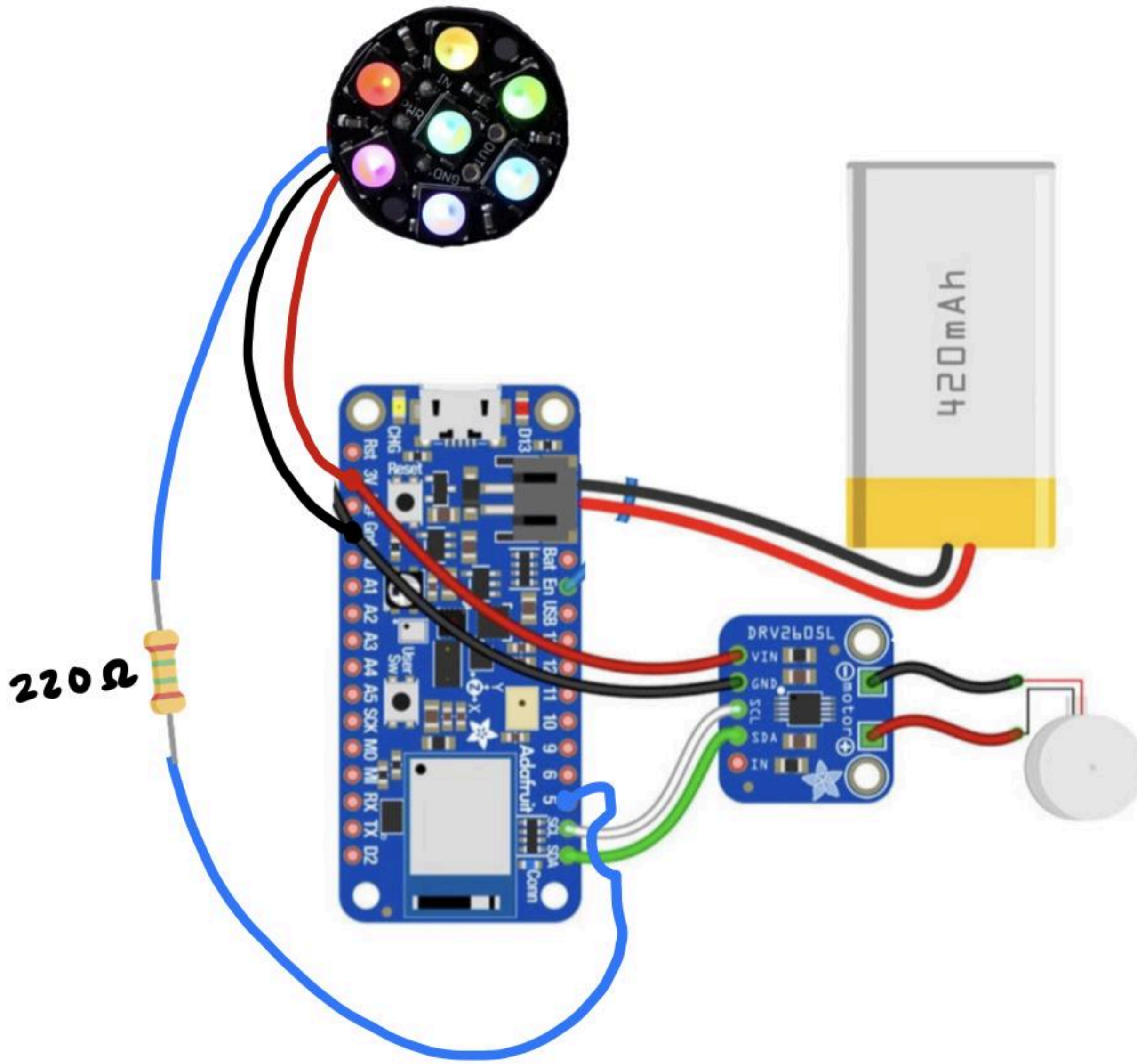
The EmotiSense system utilizes a custom-built machine learning model for emotion recognition, developed using Python and OpenCV. This model is trained on the FER-2013 dataset, which provides a diverse set of facial expressions categorized into several emotion classes. The use of OpenCV facilitates efficient image processing and real-time video analysis, crucial for the model to perform well in dynamic, real-world scenarios. The Python environment supports using the PyTorch framework to build our model. The Nvidia Jetson, a key component of our hardware setup, runs the trained model. This powerful device is capable of handling the intensive computational needs of real-time image processing and inference, ensuring low-latency responses that are necessary for the seamless operation of the EmotiSense bracelet.

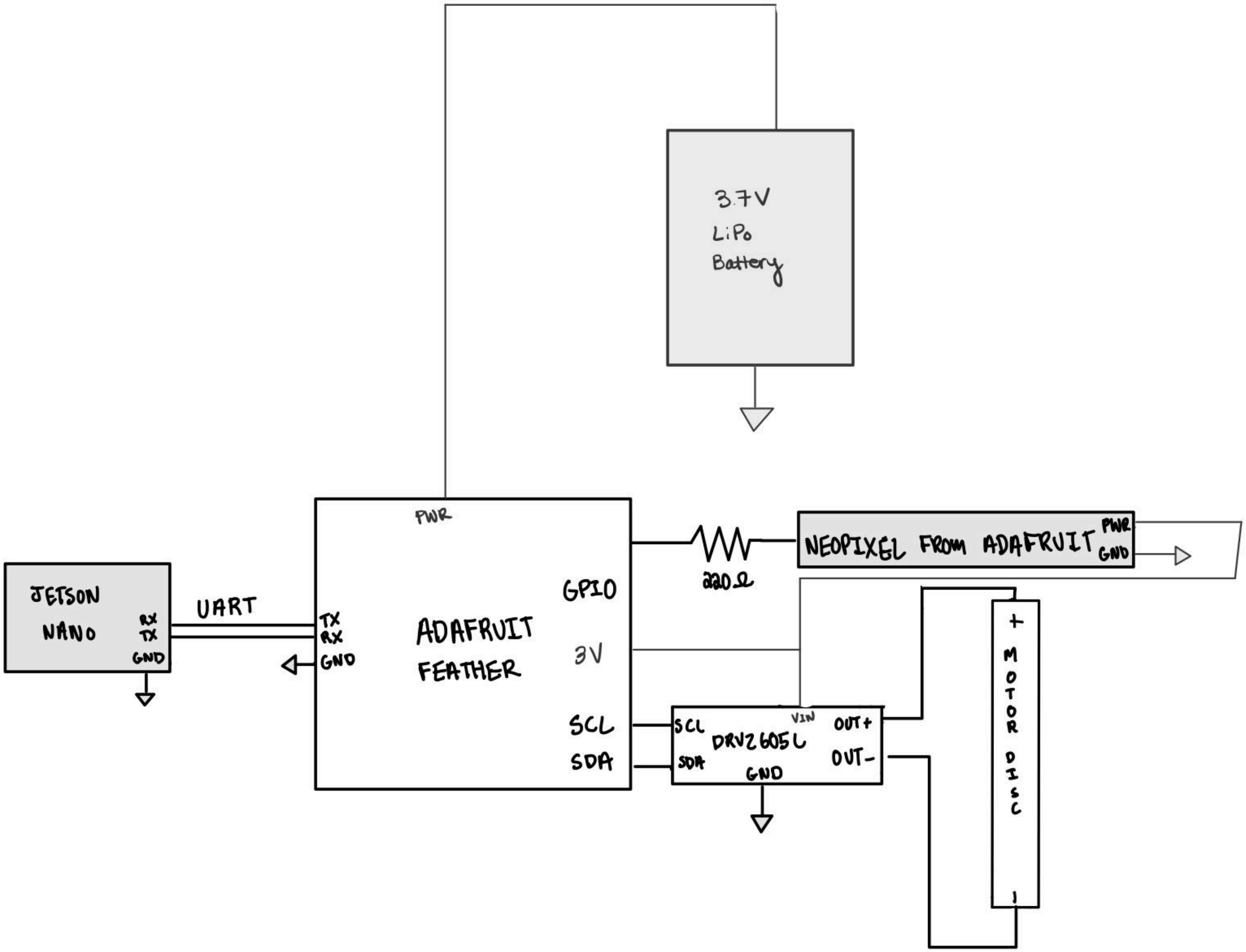
The EmotiSense web application is developed using Django, a high-level Python web framework that ensures rapid development and clean, practical design. AJAX is employed to enable asynchronous data communication between the web application and the server, ensuring that the user interface is responsive and can update in real-time without reloading the page. This is particularly important for displaying real-time emotion detection results smoothly and continuously. The web application is deployed on an Amazon EC2 instance, providing a scalable and reliable hosting environment. This allows the EmotiSense system to handle varying loads efficiently, accommodating potentially large numbers of users without degradation in performance. Communication between the web application and the Nvidia Jetson, which processes the emotion detection, is facilitated using the Python Requests package. This package is robust and simplifies HTTP requests, making it an excellent choice for server-to-server communication. API calls from the Jetson are sent to the web application over a TCP connection, ensuring reliable and ordered data transmission. This setup allows the system to maintain a continuous flow of data between the emotion recognition model and the user interface, facilitating a dynamic user experience. This comprehensive integration of hardware and software components ensures that EmotiSense delivers accurate and timely emotion recognition, enabling users to engage more effectively in social interactions.

In addressing public health, safety, and welfare considerations, the design of the EmotiSense system focuses on being non-invasive and discreet. This ensures that both neurodivergent individuals and those they interact with can use and be around the device comfortably without feeling self-conscious or uncomfortable. The discrete nature of the device plays a crucial role in fostering natural social interactions, which is essential for effective communication and emotional recognition. Furthermore, the EmotiSense system is designed with minimal environmental impact by utilizing low-power components and efficient software algorithms that reduce energy consumption. Economically, the system leverages cost-effective technology and open-source software to keep production and operational costs low, enhancing its feasibility for real-world application and widespread adoption. These considerations ensure that the device is not only effective and safe but also accessible and sustainable.

IMPLEMENTATION DIAGRAMS:




220Ω
420mAh
DRV2605L
VIN
GND
SCL
SDA
IN
motor
Adafruit


3.7V
LiPo
Battery
PWR
JETSON
NANO
RX
TX
GND
UART
TX
RX
GND
ADAFRUIT
FEATHER
GPIO
3V
SCL
SDA
220Ω
NEOPIXEL FROM ADAFRUIT
PWR
GND
SCL
SDA
VIN
DRV2605L
GND
OUT+
OUT-
+
MOTOR DISC

TEST, VERIFICATION, AND VALIDATION

Comprehensive testing and validation are conducted to ensure that the EmotiSense bracelet meets design and use-case requirements. The overall latency from emotion detection to bracelet actuation is measured using software timers and hardware tools such as an oscilloscope. The system should achieve a latency of 150 milliseconds, verifying its real-time performance. When transitioning to BLE, latency comparisons with the current UART-based implementation will be made. The BLE latency should remain within 20 ms of the UART baseline to ensure no significant impact on the system's responsiveness. User studies will assess how accurately participants can identify the emotions conveyed by the bracelet's feedback. At least 90% of participants should correctly identify the haptic and visual cues for each emotion within 3 seconds under various environmental conditions. The system's feedback patterns for different emotions will be evaluated over 20 trials, with at least 95% of these trials showing consistent feedback for the same emotion, confirming the reliability of the bracelet.The system's power consumption will be measured across different modes (e.g., maximum vibration intensity and full LED brightness vs. lower settings). The bracelet should achieve at least 4 hours of operation under typical usage conditions, which involve moderate vibration intensity and 50% LED brightness.

To validate the timing requirements for the overall system, end-to-end latency measurement will be essential. For the web application, latency will be assessed by using the latency testing functionality built into python requests total seconds elapsed functionality, which will give the time between when the requests is sent and when the requests gets acknowledged. For the embedded system (bracelet and iPad feedback), we will use cameras alongside software timing/logging to measure the exact time from emotion detection by the camera to the activation of feedback mechanisms. This test directly ties to the design requirement of feedback latency (150 ms for the bracelet and 300 ms for the iPad display) and validates the real-time interaction capability critical for seamless user communication.

To verify that the web application's UI meets the intuitive use requirement, we will conduct usability testing. The usability tests will involve a participant using the EmotiSense system, and another participant having a conversation with the user. After the conclusion of a 3 minute conversation using the system, the user will evaluate distraction, communication, and overall usability on a scale one to ten for the two output forms. The conversation participant will also evaluate the disruption that the devices caused from their point of view also on a scale of one to ten..

To verify the accuracy of the emotion classification model, two key tests will be conducted. First, we will perform automated testing against a standardized data set: using a benchmark dataset from Kaggle (FER-2013)., we will test the model's accuracy in identifying various emotions. This quantifies the model's performance and checks against the design requirement of 70-75% accuracy. Through this test, we will have a very clear picture of how effective our model is, and be able to know if it is up to our specification. The second test we will do is parallel human-model comparison testing. In this test, human participants will classify emotions from a set of video clips, and their responses will be compared against the model's outputs. This direct comparison helps validate the system's utility by ensuring that the model performs on par with unassisted human emotion recognition (for neurotypical individuals, who are strong at classifying emotions), directly supporting the use-case need for enhancing emotional understanding among neurodivergent users. We will use a set of 50 unique video clips to accomplish this goal.

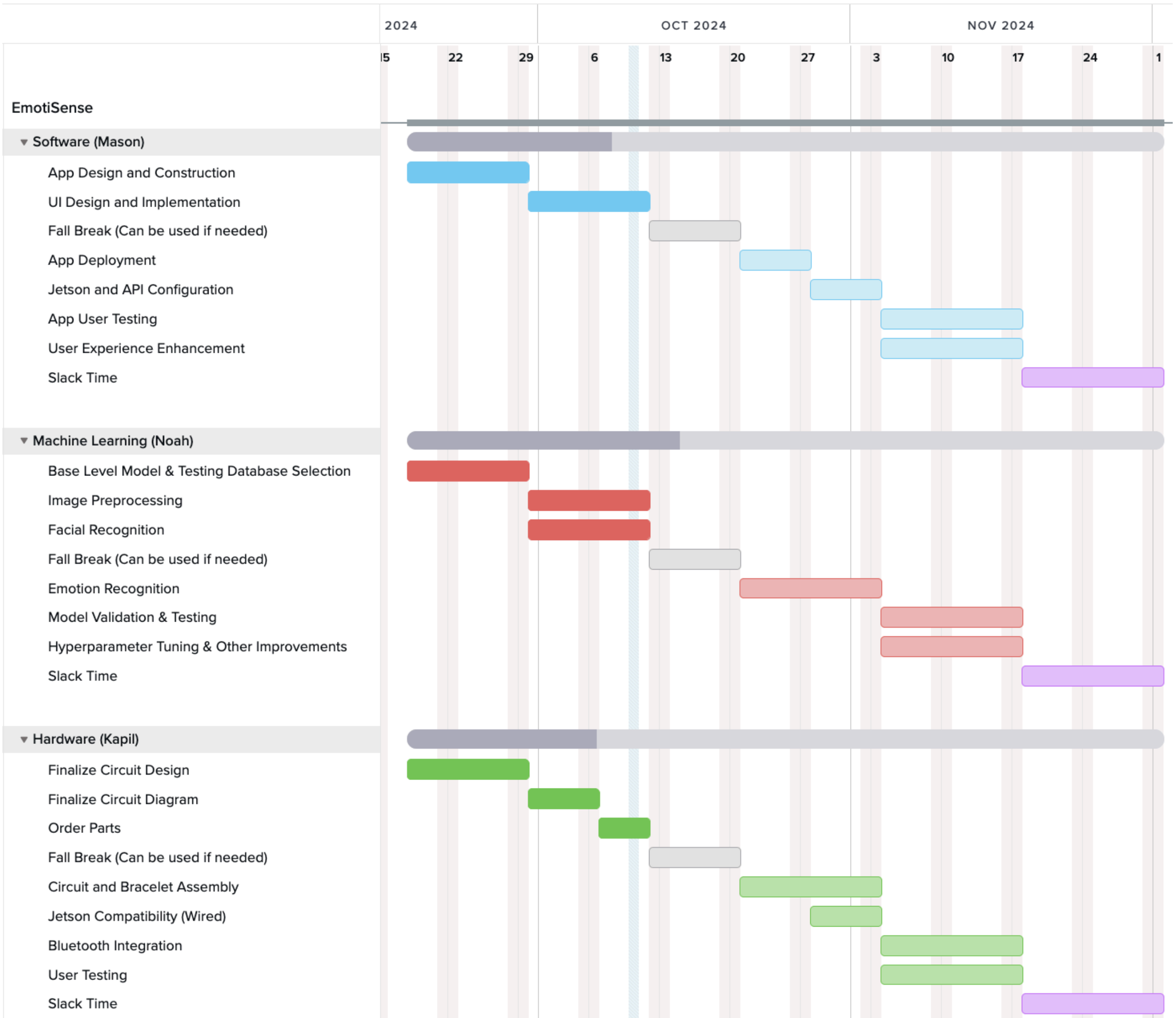
Each test is designed to critically evaluate and validate specific aspects of the EmotiSense system, ensuring that both the technical specifications and the practical use-case requirements are met effectively. This structured approach not only supports rigorous validation but also aligns closely with academic standards for empirical testing, providing a robust framework for the successful implementation of the EmotiSense product.

PROJECT MANAGEMENT

i. BILL OF MATERIALS

Component	Part/Model Number	Manufacturer	Source	Cost	Description
Jetson	Jetson AGX Xavier	Nvidia	18-500	$0	From the Class Inventory
iPad	iPad Pro 11-Inch (M4)	Apple	Mason	$0	Mason’s iPad used for Web App Display
Camera	Logitech C920x HD Pro Webcam	Logitech	Noah	$0	Noah’s webcam used for Camera/Input
Microcontroller	Assembled Adafruit HUZZAH32 – ESP32 Feather Board - with Stacking Headers	Adafruit	18-500	$21.95	Microcontroller for bracelet circuit
Haptic Motor Controller	DRV2605L Haptic Motor Controller - STEMMA QT / Qwiic	Adafruit	18-500	$7.95	Part of haptic feedback actuator in bracelet circuit
Neopixel	NeoPixel Jewel - 7 x 5050 RGBW LED w/ Integrated Drivers - Natural White - ~4500K	Adafruit	18-500	$6.95	Visual feedback actuator in bracelet circuit
Haptic Feedback Motor Disc	Vibrating Mini Motor Disc	Adafruit	18-500	$1.95	Part of haptic feedback actuator in bracelet circuit
Battery	Lithium Ion Polymer Battery with Short Cable - 3.7V 420mAh	Adafruit	18-500	$6.95	Battery for bracelet

ii. SCHEDULE AND TASK DELEGATION



The schedule has been updated to reflect the fact that A. we are no longer using a PCB for our circuit and B. that we are building our model from scratch. Our progress is up to date with our timeline as of the submission of this document.

Sources:

Barros, P., Churamani, N., & Wermter, S. (2015). Real-time neural architecture for emotional face-to-face interaction. *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*.

Zhou, Z., Cao, S., & Duan, F. (2019). Low-latency emotion recognition from facial expressions: A real-time approach. *International Conference on Computer Vision Systems (ICVS)*.

Li, S., Deng, W., & Du, J. (2018). Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high-functioning autism. *Journal of Autism and Developmental Disorders, 32(3)*, 189–194.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6(3-4)*, 169-200.

Dautenhahn, K., Werry, I., & Harwin, W. S. (2009). Robot companions for socially and emotionally interactive applications. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.

Picard, R. W., & Klein, J. (2002). Computers that recognize and respond to user emotion: Theoretical and practical implications. *Interacting with Computers, 14(2)*, 141–169.

Pentland, A. (2005). Socially aware computation and communication. *Computer, 38(3)*, 33-40.

"Accuracy of Facial Emotion Recognition (FER)". Retrieved from Morphcast.

"Real-time social perception in individuals with autism". Retrieved from Molecular Autism.

"The Emotional Consequences of Prosocial Behavior in Extraordinary Times: The Case of the COVID-19 Pandemic". Retrieved from SpringerLink.

"Understanding the Mechanisms of Social Cognition in the Context of Autism". Retrieved from Frontiers.