Final Report for Computer Systems Lab

Voca: Voice Activated Non-Invasive Prosthetic Arm

Anish Suvarna, Aarav Gupta, Srikar Kode

5/28/2025

Dr. Yilmaz - Period 3

Table of Contents

Abstract

This project presents Voca, a non-invasive, voice-activated prosthetic arm designed to provide affordable, intuitive, and emotional resonant solutions for individuals with upper limb loss. Traditional prosthetics are often prohibitively expensive, uncomfortable due to invasive control methods, and emotionally detached from their users. Voca addresses these challenges through the integration of natural language processing for seamless voice control, AI-based autonomous grasping for functional versatility, and a 3D-printed, customizable design that prioritizes comfort and accessibility, all under a $1500 budget. The system uses a microcontroller, depth camera, and tactile sensors to enable multiarticulation, haptic feedback, and realtime object detection. Results from prototype testing demonstrated successful grasping and user-directed actuation in varied conditions. Voca reimagines prosthetic technology not only as a tool for mobility, but as meaningful extensions of a person’s will, laying the foundation for a more compassionate future in assistive technology."

I. **Introduction**

Around the world, more than 22 million people live with limb loss, and for many of them, getting a prosthetic arm isn't easy. Most high-quality prosthetics cost anywhere from $15,000 to $80,000, which just isn't realistic for a lot of people. On top of that, the ones that are available often rely on invasive sensors that can cause skin irritation or discomfort. They're also not always easy to use or connect with on a personal level, which leaves users feeling like they're using a tool rather than something that's really part of them.

When we first started this project, we were struck by how little focus there was on accessibility, comfort, and emotional connection in prosthetic design. It felt like the people who actually use these devices were being left behind. That's what drove us to build Voca, a prosthetic arm that's not only functional and easy to use, but also affordable and non-invasive.

Voca is a voice-activated, AI-powered prosthetic that can be built for under $1,500, a huge drop from the price of traditional options. We used 3D printing, natural language processing, and machine learning to create a system that listens to your voice and responds in real time. It doesn't need skin-contact sensors or complex manual controls. Just speak, and it moves.

What really sets Voca apart is how it combines practical function with emotional design. It offers haptic feedback, a sleek carbon fiber look to reduce stigma, and a voice assistant to make interactions feel more natural. At the end of the day, our goal wasn't just to build a cheaper prosthetic, it was to give people a tool that helps them feel confident, in control, and human.

II. Background

When we looked into existing prosthetic solutions, we found that most of the major players, like Ottobock, Taska, Ossur, and Covvi, focus heavily on high-tech functionality, but with a serious trade-off: price and accessibility. These devices often include advanced EMG (electromyographic) sensors that detect muscle signals through the skin or implanted electrodes. While this allows for responsive movement, it also means users deal with skin irritation, fatigue, and in many cases, a steep learning curve just to operate the device comfortably. And again, the price point for these systems often ranges between $15,000 and $80,000, which puts them out of reach for many who actually need them.

On the software side, most EMG-based systems rely on proprietary algorithms and require specialized training to operate. There's very little focus on natural interfaces like voice commands or non-invasive control, and even less effort to build prosthetics that feel emotionally engaging.

That's where we saw an opportunity. We wanted to eliminate the need for any invasive sensors entirely. Instead, we turned to Natural Language Processing (NLP) and machine learning to create a prosthetic that responds to voice input. Using an onboard microphone and a speech-to-text engine, our system processes user commands and maps them directly to movements via servos. For grasping objects, we integrated a depth camera and an AI model based on a generative residual convolutional neural network. This model predicts grasping poses based on real-time video, allowing the prosthetic to locate and interact with objects, even in cluttered environments.

III. **Applications**

Voca is designed to help amputees regain independence by offering an affordable, intuitive, and non-invasive prosthetic arm. Its versatility makes it useful across a wide range of real-world scenarios. For everyday users, Voca enables people to perform common tasks, like picking up objects, opening doors, or using utensils, using only their voice. This makes it especially helpful for users with limited strength or dexterity, as there's no need for physical sensors or complex control systems. In medical and rehabilitation settings, Voca can serve as a training or transitional device, allowing patients to build confidence and motor control without the discomfort of invasive EMG systems. It’s also a strong option for veterans or trauma victims who want a functional, responsive arm without the steep price tag. Beyond personal use, Voca has the potential to make a global impact. Its low cost and modular design mean it could be distributed through humanitarian aid organizations or public health programs in underserved regions, where access to advanced medical technology is limited. Finally, because it’s built using open-source hardware and software, Voca can also be used in educational settings as a platform for learning about robotics, machine learning, and human-centered design. Whether in a home, hospital, or classroom, Voca offers a practical and empowering solution to the challenges faced by people with limb loss.

IV. Materials

To bring Voca to life, we combined a range of affordable yet capable hardware components with open-source software tools, all tied together through iterative design and hands-on prototyping. A central part of the system was the Arduino Uno, which acted as our main control board. It handled communication between different modules, processed input commands, and controlled output to the motors. For motion, we used a set of high-torque servo motors to actuate the fingers and wrist. These servos were chosen for their reliability, speed, and compatibility with our mechanical linkage systems.

To give the hand lifelike movement, we designed a 4-bar linkage system for each finger, allowing for coordinated, smooth articulation. For the wrist, we implemented a spherical parallel manipulator that enabled multi-directional rotation, mimicking natural wrist motion. Finger movement was planned and stabilized using PID control loops and inverse kinematics, ensuring accuracy even with limited hardware.

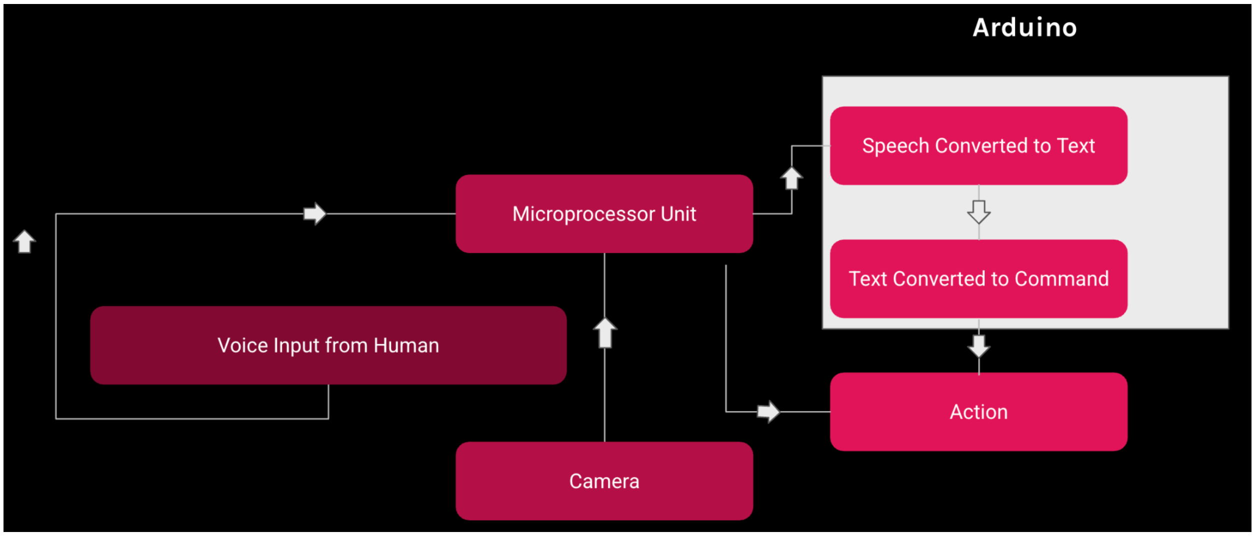
A major feature of Voca was its ability to recognize and interact with objects autonomously. For this, we used an Intel RealSense depth camera, which captured real-time 3D spatial data of the environment. The camera provided both depth and RGB information, which was processed by a generative residual convolutional neural network that we trained to identify objects and predict optimal grasping positions (6D pose estimation). This AI grasping model was coded in Python using machine learning libraries, and communicated with the Arduino to drive the appropriate servo motions.

For voice control, we used a microphone module and developed a custom natural language processing (NLP) system. This system interpreted spoken commands like “grasp” or “release,” converted them into executable actions, and triggered motor control via serial communication. The NLP backend ran on a laptop using Python and open-source speech-to-text libraries.

All mechanical parts were created using 3D printing, which allowed us to rapidly iterate and customize the hand’s design. We used PLA for initial prototypes and switched to carbon fiber-infused filament for the final build to improve strength, reduce weight, and create a sleeker appearance. We also integrated tactile sensors for basic haptic feedback and a small audio module to play confirmation sounds when commands were processed successfully.

For design and simulation, we used Fusion 360 to model the prosthetic components, run mechanical stress analyses, and ensure smooth motion before printing. Code for motion control was written in C++ (Arduino IDE), while all AI and NLP logic was done in Python.

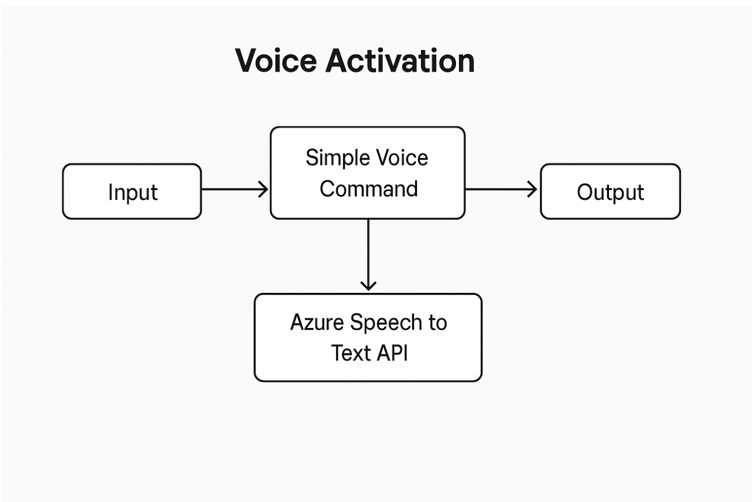
V. **Methods**



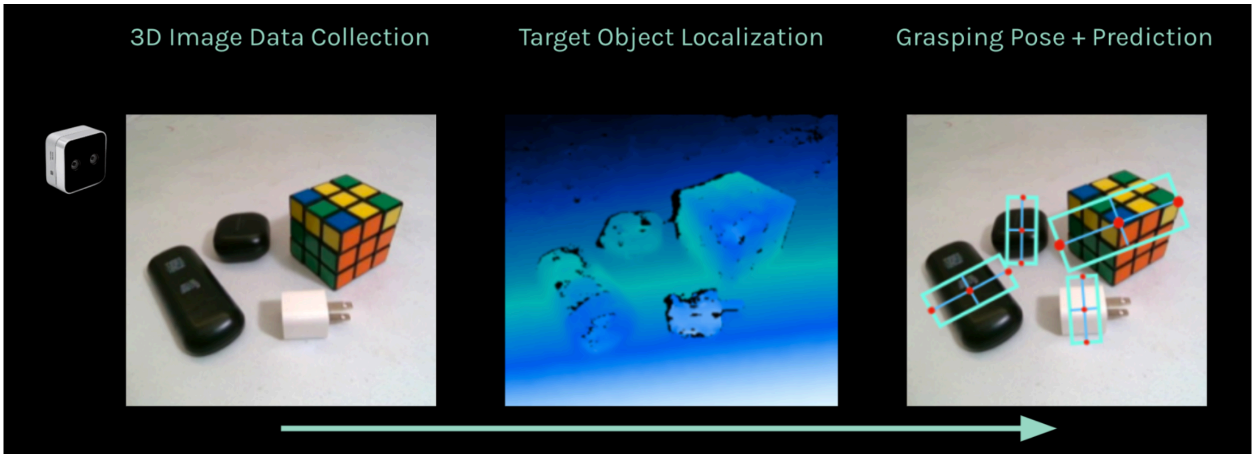
The VOCA prosthetic system is a combination of depth perception, voice control, and mechanical engineering designed to provide amputees and upper-limb-differently-abled individuals with a natural, fluid way of dealing with their environment. VOCA does away with the computationally intensive need and instead chooses a deterministic, hardware-oriented pipeline that combines real-time voice recognition and depth sensing for facilitating responsive grasping. This allowed the system to be light-weighted, reliable, and appropriate to embedded applications while maintaining high mechanical and behavioral realism.

VOCA's speech command system recognition is powered by Microsoft Azure's Speech-to-Text API. We chose this solution because it has low latency, is cross-platform, and has a high transcription accuracy within controlled environments. Instead of building a full natural language processing pipeline, we restricted the command vocabulary to exact, high-confidence terms such as "grab," "release," "stop," and "reset." These were executed in real-time on a Arduino Mega that parsed Azure's JSON response and filtered on confidence scores above 85%.

Commands were only accepted when speech detection and object proximity conditions were both met, which minimized false positives. We initially attempted to include offline models such as Vosk and DeepSpeech so that the prosthetic would be able to work without internet dependency; however, we found that latency and word error rates were significantly higher on embedded hardware, resulting in missed commands and inconsistent performance. Azure ultimately offered the best trade-off between reliability and ease of deployment, especially in environments where Wi-Fi is present.



To complement the speech interface, VOCA uses an Intel RealSense D435 depth camera positioned above the wrist joint to keep track of the object distance in real time. We used its depth stream to compute the Z-axis distance between the palm's centroid and possible target objects. The grasp routine was invoked only when an object entered a pre-defined bounding box, 5 to 15 centimeters in front of the palm, and remained there for three successive frames in order to ensure temporal coherence. This was a form of spatial gating to avoid accidental triggering by movement at rest or noisy sensor readings. We also experimented with an ultrasonic proximity sensor in the early stages for its simplicity and low power usage, but its limited beam angle and confined field-of-view rendered it unreliable for detection of objects sideways, especially during rotation of the wrist. The RealSense system not only provided superior depth granularity but also stereo-based robustness to varying light conditions in the environment.



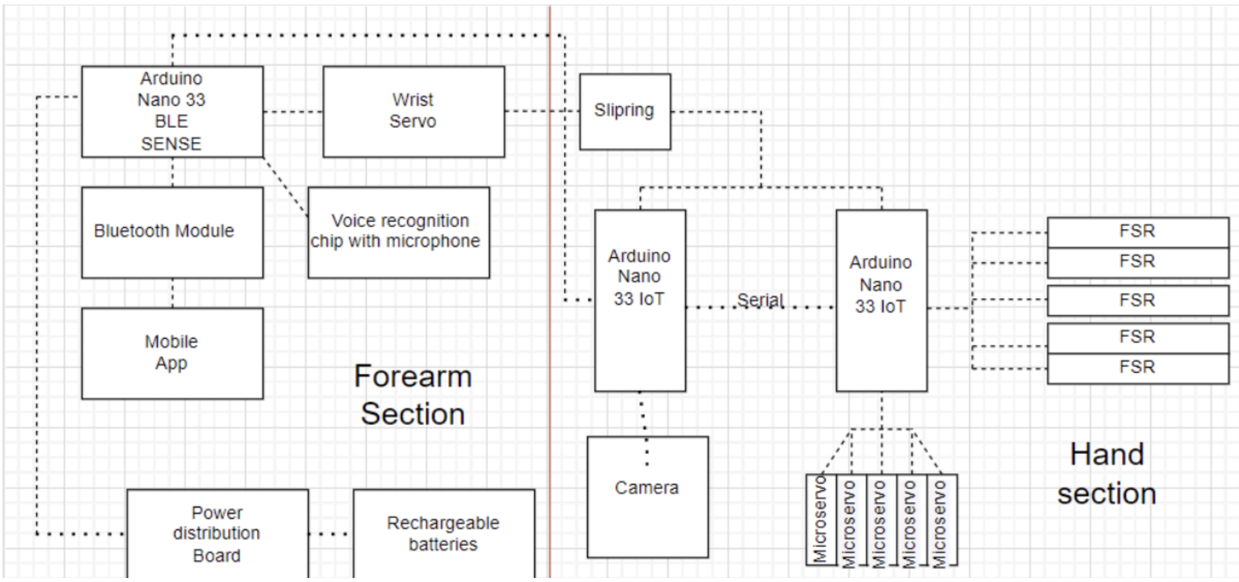
Mechanically, VOCA's hand is designed for bio-inspired motion fidelity. Its finger is constructed from a 3-bar planar linkage mechanism, optimized to mimic the curling motion of human phalanges while reducing rotational stress at joints. The mechanical advantage of this linkage mechanism allowed VOCA to use lightweight servos without affecting grip force. The wrist employs a spherical parallel manipulator (SPM), with three rotational degrees of freedom, pitch, yaw, and roll, and is required for dynamically orienting the hand towards objects of different shapes or orientations. The SPM configuration also achieves compactness and torque gain, both of which are ideal for a prosthetic device where force and size are of equal concern. These movements are controlled by an in-house inverse kinematics solver, which computes joint angles required to reach some desired end-effector pose without exceeding joint limits or causing collisions with self.

All components of VOCA were 3D-printed with a modular structure. Carbon-fiber-infused PLA was used to produce rigid structural parts, while TPU produced flexible finger pads and areas of tension. Bowden cable sleeves guided the tendon system, and Kevlar threads were used for their strength-to-weight ratio. Each servo motor was placed in a forearm cavity and powered by a dedicated PWM driver board. While this mechanical design served adequately in most uses, we did experience tendon slipping when used repeatedly, especially in successive quick grasps. To mitigate against it, we attempted to modify with custom tensioners and screw-cable locks, but they introduced bulk into the design that was not acceptable. Were we given more time, we would have explored magnetic clutches or feedback motor braking systems to provide greater tension control and grip modulation.

VOCA is energized by an internal LiPo battery pack with current-limiting regulator and thermal cutout circuitry. An Arduino is used as the main controller for Azure API calls, interpreting sensor data, and transmitting actuation commands to a secondary microcontroller (Arduino Mega) that handles real-time motor control. In earlier releases, we attempted to utilize a single microcontroller to power the entire system, but quickly ran into memory issues and timing issues when trying to process voice input as well as camera input simultaneously. Separation of high-level control logic and low-level control logic into a two-processor design allowed cleaner division of labor and much better responsiveness of the system.

Reflecting on what else we could have done, one large opportunity would have been employing real-time sensory feedback such as pressure sensors or current-sensing resistors in order to feel the force being applied in a grasp. Without this feedback loop, VOCA is in open-loop control, meaning it cannot modulate grip strength in relation to object fragility or user desire. Another upgrade would be to replace Azure's cloud architecture with a more advanced offline speech recognition system if we had resources and time to make it more streamlined for embedded deployment, to maintain the device function perfectly offline. We would also have explored multi-material printing for compliant integrated structures to reduce the number of parts and improve comfort and mechanical robustness had we been able to stretch our manufacturing budget and available time.

In hindsight, if we were given the chance to work on the project again, we would have prioritized user-centered testing much earlier in development. We did most of our testing as lab bench setups and simulated conditions, maximizing mechanical and software performance but limiting our understanding of how real users would use VOCA in dynamic, unscripted environments. By integrating real-time user feedback from target users sooner in the prototyping cycle, we could have made voice commands, grasp profiles, and mechanical ergonomics more inclusive and intuitive. Lastly, a formal data-logging system would have allowed us to track failure rates, servo loads, and command recognition accuracy over time, govern future improvements with empirical evidence.



VI. **Results**

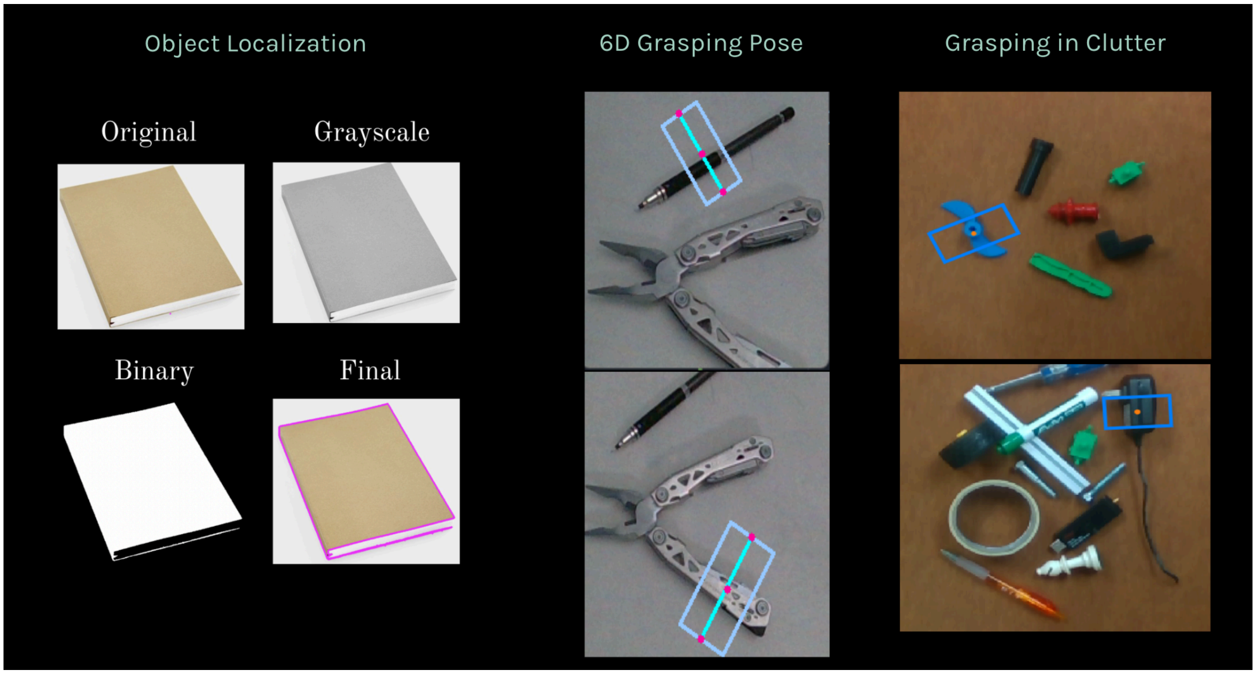
By the end of the project, we had successfully created and tested multiple working prototypes of the Voca prosthetic hand. Each version brought us closer to a system that was more responsive, more capable, and more natural to use. One of our biggest milestones was implementing reliable "grasp" and "release" functions through voice control. Using a natural language processing algorithm we developed, Voca is able to take a user’s spoken command, convert it into text, interpret the intent, and trigger the appropriate physical action through our servo-based control system.

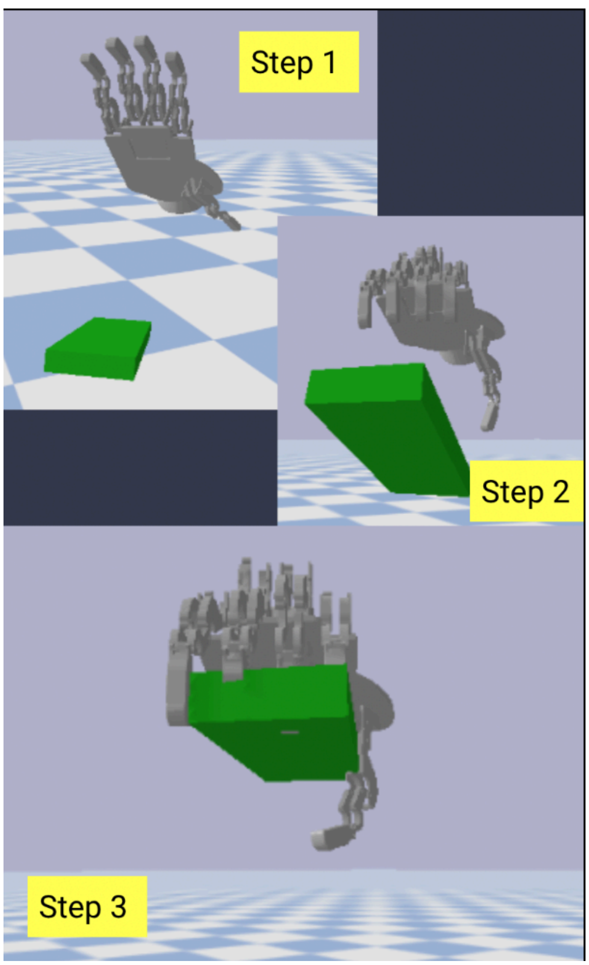
To support this, we wrote custom Arduino-compatible code that interfaces with our servo motors and maps each interpreted command to specific joint movements. The code runs efficiently on a low-power microcontroller, enabling real-time motion based entirely on verbal input. This made the interaction feel smooth and hands-free, exactly what we envisioned when we started building a non-invasive interface.

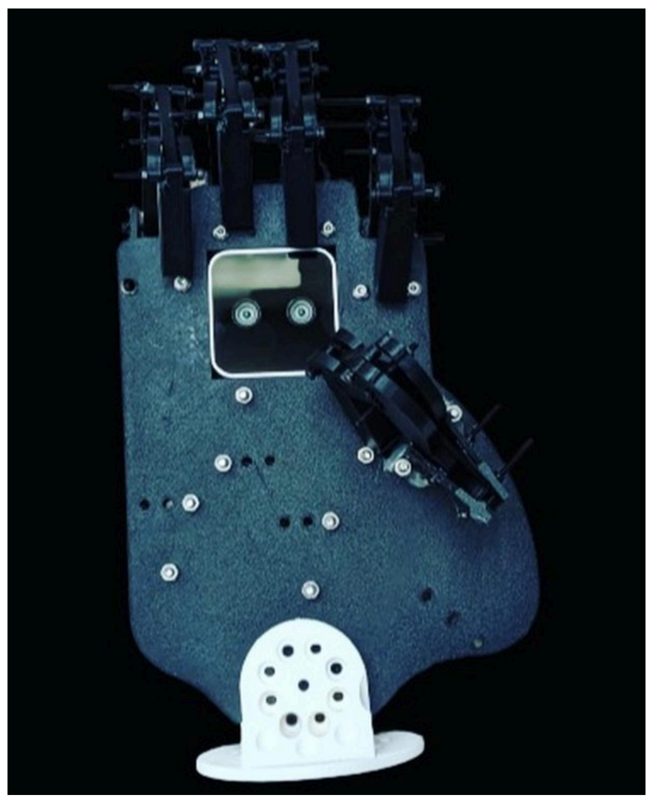
We also developed an AI-based object detection and grasping module that uses a depth-sensing camera to locate objects in 3D space. The software, powered by a generative residual convolutional neural network, analyzes the scene, estimates a 6D grasp pose (position and orientation), and sends that data to the microcontroller to trigger a corresponding movement. This allowed Voca to recognize and interact with various objects, even in cluttered environments. The system was especially effective at handling medium-sized objects like blocks, bottles, and utensils.

Mechanically, we iterated through several designs to improve hand articulation. Our final prototype used a 4-bar linkage system to coordinate finger motion, supported by PID control loops that helped stabilize the movement. The wrist was built around a spherical parallel manipulator, giving it a wide range of motion while maintaining strength and balance. For finger positioning and planning, we used inverse kinematics algorithms to calculate realistic, human-like motion paths.

On the feedback side, we integrated basic haptic sensors and audio signals into the design. Although these systems were simple, they helped users confirm when an action, such as a grasp, was successfully completed. These features, along with the voice assistant functionality, gave the arm a more interactive and engaging feel.


Object Localization
6D Grasping Pose
Grasping in Clutter
Original
Grayscale
Binary
Final


Step 1
Step 2
Step 3



VII. **Limitations**

Even though Voca is a big step forward in making prosthetics more affordable, intelligent, and user-friendly, we know there's still a lot of work to do. One of the biggest limitations right now is the lack of EEG-based control. Because the arm doesn't connect directly to the user's brain signals, we rely entirely on voice commands for input. While voice activation works well in many situations, it isn't always ideal, especially in loud environments, or for users who have trouble speaking clearly, or who speak in a way the system doesn't recognize. EEG would allow for much smoother, more intuitive control, where you could move the arm just by thinking about it, but that kind of technology is expensive and complicated to implement, especially at the low cost we're aiming for. Still, it's something we really want to explore as we continue developing Voca.

Another trade-off we've had to make is due to our budget-conscious design. We built Voca to stay under $1,500, which is a huge improvement compared to most commercial prosthetics, but it also means we're working with cheaper microcontrollers and limited onboard processing power. That restricts how fast and precisely we can run things like voice recognition or AI-based grasping. The system performs well for basic tasks, but it can struggle with more detailed movements or when trying to handle multiple functions at once, like listening to commands and identifying an object at the same time.

There are also some limitations when it comes to the physical design. While our use of 3D printing and carbon fiber makes the arm light and relatively durable, it doesn't feel like a real human arm. It's missing the warmth, softness, and flexibility of skin and muscle. That can make the prosthetic feel a little mechanical or disconnected, and it may affect how users emotionally relate to it. This isn’t just about comfort, it’s about helping people feel like the prosthetic is truly part of them.

Lastly, while we do have basic haptic and audio feedback, it’s not as advanced as what you'd find in high-end devices. Voca can’t sense how firmly it's gripping an object, and it doesn’t adjust its movement based on texture or resistance. That kind of feedback loop would be incredibly helpful for making the arm feel more responsive and lifelike.

VIII. Conclusion

Voca began as an ambitious idea: could we create a smart, non-invasive prosthetic arm that was actually affordable, and meaningful, for real people. By the end of this journey, we proved that it’s not only possible, but absolutely necessary. Through months of designing, coding, testing, and refining, we developed a fully functional prosthetic hand capable of responding to voice commands, performing autonomous object grasping, and offering a degree of emotional connection through haptic and audio feedback, all for under $1,500.

What sets Voca apart isn’t just the technology, but the philosophy behind it. We didn’t want to just copy existing systems with fewer features, we wanted to rethink what prosthetics could be if accessibility, empathy, and intuitive design were the priorities from day one. By using natural language processing, a depth camera, machine learning models, and a custom-built mechanical system, we created a device that doesn’t rely on invasive sensors or overwhelming complexity. Instead, it speaks the user’s language, literally.

Of course, we know Voca is still a work in progress. It doesn’t yet have EEG integration, clinical certification, or highly realistic material design. But what we’ve built is a strong foundation: a system that works, that's modular, that's open to growth, and that makes prosthetic technology feel more human.

More than anything, this project taught us the power of interdisciplinary thinking, where engineering, AI, design, and compassion intersect. Voca isn't just a robotic hand. It's a small glimpse of what the future of assistive tech could look like: affordable, intelligent, and made with real people in mind.

IX. Future Work

Looking ahead, we view Voca not just as a capstone project, but as the foundation of a long-term vision to revolutionize prosthetic technology through accessibility, intelligence, and human-centered design. The core challenge we set out to address, creating a functional, affordable, and emotionally resonant prosthetic arm, has only revealed further layers of complexity and opportunity. Our next steps involve transitioning Voca from a promising prototype into a clinically validated, widely distributable device that can impact real lives.

A key part of this progression is deepening our research and expanding our understanding of both the technical and human aspects of prosthetic use. To that end, we are actively exploring the possibility of partnering with cutting-edge research institutions, especially Stanford University's Biomechatronics and Biorobotics Lab, known for its interdisciplinary work at the intersection of prosthetics, neuroscience, and machine learning. Through such a collaboration, we hope to refine Voca's control systems, improve our AI-based grasping capabilities, and explore new methods for voice interpretation that adapt to individual users over time. Moreover, we aim to study long-term user adaptation, comfort, and cognitive load through structured clinical observations and trials, insights that are crucial for real-world deployment.

To reach this next stage, increased funding will be essential. By working with academic partners, we intend to apply for federal research grants such as those offered by the National Science Foundation (NSF) and National Institutes of Health (NIH), alongside private foundation support for assistive technologies. These resources would support additional prototyping, large-scale user testing, and the rigorous documentation required for regulatory submissions.

Once we've iterated on our design with clinical feedback and produced comprehensive safety and performance data, we plan to formally pitch Voca to the U.S. Food and Drug Administration (FDA) for Class II medical device approval. Navigating this process will involve submitting documentation for safety testing, risk mitigation strategies, and real-world efficacy, ensuring that Voca meets the standards required for patient use and insurance reimbursement. Successfully obtaining FDA approval would be a transformative step, allowing the prosthetic to be formally prescribed, covered by insurance, and integrated into standard medical workflows.

Following approval, our goal is to distribute Voca through a clinic-based deployment model. We envision a system where patients can be referred by physicians, evaluated by prosthetists, and trained by rehabilitation specialists to use the device effectively. Clinics would also serve as feedback hubs for continued iteration and personalized upgrades. Over time, this structure could be scaled globally, with localized manufacturing centers using open-source CAD files and voice models adapted to different languages and cultural needs. This model ensures that Voca can be delivered responsibly and equitably, without compromising on quality or user support.

References

Chen, Jie. IEEE Xplore, The Institute of Electrical and Electronics Engineers, 28 Mar. 2023, ieeexplore.ieee.org/document/10984809/.

Herrmann, Achim Jan, and Michelle Gehringer. "Open-Source 3D-Printed Upper-Limb Prosthetics: A Review of Designs, Materials, and Future Directions." PennEHRS, Oct. 2021, ehrs.upenn.edu/health-safety/lab-safety/safety-alerts/anaerobic-chamber-explosion.

Ki, Wing-Hung. "Non-Invasive EEG-Controlled Prosthetic Arm for Disabled PersonsNon-Invasive EEG-Controlled Prosthetic Arm for Disabled Persons." IEEE Xplore, The Institute of Electrical and Electronics Engineers, 1 May 2020, ieeexplore.ieee.org/document/9084122/.

Mathur, Vijay, et al. "Toward Next-Generation Non-Invasive Prosthetics: Combining AI, Voice, and Sensor Feedback Source: NCBI (Journal of NeuroEngineering and Rehabilitation)." Oman Journal of Ophthalmology, U.S. National Library of Medicine, 2 Nov. 2020, pmc.ncbi.nlm.nih.gov/articles/PMC7852412/.

https://github.com/anishsuvarna11/Voca2.0

APPENDIX

I. **CODE:**

https://drive.google.com/file/d/1HuczKu-iksCga0KooMyaemzM7AVU86J-/view?usp=drive_link