Grenze International Journal of Engineering and Technology, Jan Issue

Efficient on Device Object Detection for Prosthetic Arm

TM Geethanjali[1], Akshay Kumar[2], Nishant Nayan[3], MD Aman Alam[4] and Samarth U.S Katti[5]
[1-5]PES College of Engineering/Department of Information Science, Mandya, India
Email: geethanjalitm@pesce.ac.in, {akshaygautam010, nishant2658, alammdaman2019, samarthkatti}@gmail.com

***Abstract*— This research paper presents a novel project aimed at enhancing the functionality of prosthetic arms through the integration of efficient on-device object detection. The project leverages a Raspberry Pi, equipped with a Mobile net computer vision model, and a camera embedded within the palm of the prosthetic arm. The primary objective is to enable real-time detection of everyday objects within the arm's field of view. Upon object recognition, the prosthetic arm performs either a grip or grasp action, enhancing the user's interaction with the environment. This paper outlines the technical details, implementation process, and the potential impact of this innovative system on the lives of individuals with limb loss, offering increased autonomy and improved quality of life.**

I. INTRODUCTION

In the realm of assistive technology, prosthetic arms have played a pivotal role in restoring mobility and independence to individuals with limb loss. Over the years, technological advancements have significantly improved the design and functionality of prosthetic limbs, offering users greater control and versatility. However, one critical aspect that has remained a challenge is the seamless interaction of prosthetic arms with the surrounding environment. This research introduces a project focused on addressing this limitation—Efficient On-Device Object Detection for Prosthetic Arm. The core objective of this project is to empower individuals with prosthetic arms to better engage with their daily surroundings by harnessing the capabilities of computer vision. By integrating a Mobilenet computer vision model into a Raspberry Pi platform and embedding a camera within the palm of the prosthetic arm, real-time object detection becomes a reality. The significance of this innovation lies in its potential to bridge the gap between prosthetic technology and the complex, dynamic world in which users operate. The ability to identify and respond to everyday objects in real time opens doors to a more intuitive and functional user experience. Whether it's grasping a cup of coffee, picking up a smartphone, or manipulating tools, this technology has the potential to redefine the capabilities of prosthetic arms. This paper unfolds the technical intricacies of our project, detailing the design, implementation, and performance of our Efficient On-Device Object Detection system. Furthermore, it explores the broader implications of this technology, ranging from improved user autonomy to enhanced quality of life. As we delve into the inner workings of this innovation, we invite readers to join us on this journey toward a more inclusive and technologically empowered future for individuals with limb loss.

II. RELATED WORK

Advancements in prosthetic technology have significantly enhanced prosthetic functionality. In the work by Scott et al. [1], they explore the use of sensors attached to amputee stumps for prosthetic control, underlining the pivotal role of sensor technology and machine learning in augmenting prosthetic capabilities. Likewise, Weiner et al. [4]

Grenze ID: 01.GIJET.10.1.339_1

delve into semi-autonomous prostheses for daily activities, emphasizing the significance of sensor data and its impact on intuitive grasp control. These studies collectively emphasize the transformative potential of sensor technology and machine learning in elevating prosthetic functionality.

Limitation: These cited papers overlook a critical aspect—the power consumption of the sensor network required to control the prosthetic. The integration of power management within the system becomes essential, albeit adding complexity.

There exists a diversity of prosthetic types, each offering distinct advantages and limitations. Schweitzer et al. [2] argue in favor of body-powered prostheses for specific occupational contexts, while Zuo et al. [3] advocate for myoelectric prostheses, which rely on external power and necessitate periodic recharging. Seitz et al. [8] provide an exhaustive review of control methods employed in prosthetic hand grasping, encompassing myoelectric control, body-powered control, and hybrid approaches, carefully elucidating their respective strengths and weaknesses. Meanwhile, Castellini et al. [12] conduct a comprehensive survey of both invasive and non-invasive interfaces for prosthetic limb control, encompassing myoelectric, brain-computer interface, and peripheral nerve interface technologies. They also explore the potential for combining these interfaces to enhance prosthetic control.

Limitation: While these studies offer deep insights into control systems, they do not address the financial implications of implementing such systems without compromising the pledged functionalities.

Sensory feedback assumes critical importance in the domain of prosthetic functionality. Carrozza et al. [5] advocate for the development of upper limb prostheses that closely replicate the sensory-motor capabilities of natural limbs. Farina et al. [6] undertake the task of designing a prosthetic hand equipped with sensory feedback capabilities to optimize grasping and grasping force control. Puskar et al. [7] furnish a comprehensive overview of contemporary prosthetic technology tailored for upper limb amputations, shedding light on advancements in control methodologies, material innovations, and sensory feedback systems. They also acknowledge the existing challenges and lay the groundwork for future research directions in the field.

Limitation: Although these studies discuss the achievement of advanced sensory-motor control, they omit an essential aspect—the driving system required to harmonize with such sophisticated sensory-motor capabilities.

The integration of cutting-edge technologies, such as 3D printing and virtual reality, has ushered in transformative improvements in prosthetic design and rehabilitation. Lake et al. [9] meticulously evaluate the utilization of additive manufacturing techniques, particularly 3D printing, in the conceptualization and fabrication of prosthetic limbs. They underscore the manifold advantages of this technology, including bespoke prosthetic designs, expedited manufacturing processes, and cost-effectiveness.

Limitation: Notably, this method's inherent challenges, particularly the structural strength and vulnerability to elevated temperatures caused by electronic operations, remain unaddressed.

Huber et al. [13] venture into the realm of virtual reality technology for prosthetic training and rehabilitation. They elucidate the potential benefits, encompassing heightened patient engagement and motivation, while also acknowledging the inherent challenges when integrating this technology into prosthetic training. Skidmore et al. [10] conduct a comprehensive investigation into the impact of prosthetic use on the quality of life and functional outcomes for individuals with upper limb amputations. Their work highlights the pivotal role of prosthetic rehabilitation in optimizing the advantages of prosthetic use. Additionally, Sosa et al. [11] delve into the economic implications of prosthetic technology, meticulously examining the costs associated with prosthetic utilization and the broader economic effects on individuals, families, and society.

Limitation: These referenced studies concentrate predominantly on user interactions and cosmetic aspects rather than the development of a low-cost, low-complexity prosthetic arm with basic functionality.

III. Methodologies

The current state of prosthetic arms in India is characterized by two main limitations: cost and functionality. Imported prosthetic arms are often prohibitively expensive, making them inaccessible to a large portion of the population. On the other hand, locally available prosthetic arms in India, such as the "Inali," which is considered the most advanced in the country, employ a gesture-based control system.

Gesture-based control systems require users to perform specific gestures to execute basic actions with the prosthetic arm. While this approach may seem intuitive, it has inherent drawbacks. Firstly, gesture-based systems can be prone to failure. The accuracy of recognizing and interpreting gestures can vary, leading to inconsistent performance. This unreliability can be frustrating for users who rely on their prosthetic arms for daily activities. Gesture-based control systems can be harder to use compared to more advanced alternatives. Learning and remembering a multitude of gestures for different actions can be challenging and time-consuming for individuals with limb loss. It requires significant training and practice to become proficient in operating the prosthetic arm effectively. These typically support only a limited number of predefined actions. Users may find themselves constrained by the limited functionality and lack of adaptability. This limitation restricts the range of activities and tasks that can be performed with the prosthetic arm, potentially hindering individuals' independence and quality of life.

A. Proposed System

The proposed system uses a combination of ML and CV. A camera is placed in the palm of the artificial limb to detect the type of object that needs to be held/utilized, and based on the object, a corresponding grasp action is used, decided using Machine learning. Grasp actions are divided into two categories: power and precision. Power grasp action is used to hold larger objects, such as a door knob/handle or a bottle of water. Precision grasp is used for smaller objects or parts of a larger object, like opening a bottle cap or holding a coin. A single-board computer inside the prosthetic arm runs a TfLite model for quick object identification. Since we cannot train a model onboard, we use transfer learning supported by Tensor Flow to train a “Mobile net” model on the “COCO” dataset. "Mobile net" is chosen for efficiency and speed, crucial for a small-sized computer. The “COCO (Common Objects in Context)” dataset, with 200,000 images and around 90 object categories, labeled by Microsoft and Facebook researchers, is used. We employ a pre-trained model from Tensor Flow, tuning it for our use case, recognizing only six common daily life objects.

B. Advantages

- *Cost-Effective*: The advanced system is cost-effective, increasing accessibility. Efficient manufacturing processes and affordable components make it an economical solution compared to traditional prosthetic systems.
- *Fast Performance*: The prosthetic arm delivers rapid and responsive performance, minimizing delay between user actions and the arm's response for smooth and efficient movements.
- *Power Efficiency:* The system is engineered for power efficiency, extending battery life or reducing power consumption. Optimization of hardware and software components maximizes power utilization, reducing the need for frequent recharging or battery replacement.
- *Lightweight Design*: Prioritizing a lightweight design, the prosthetic arm ensures user comfort and minimizes strain. Lightweight materials and efficient construction techniques allow natural movements without discomfort during extended use.
- *User-Friendly*: The system features a user-friendly, easy-to-operate design with a simplified interface and intuitive controls, reducing the learning curve for users. Its limited complexity enables quick adaptation and seamless integration into daily life.

C. Database Module

For the model training, the first step is to find a reliable dataset that features most of the daily life objects that a human might encounter while using the prosthetic arm, such as water bottle, cell phone, glass etc. We chose COCO dataset (Common Objects in Context) [15] which is created by researchers at Facebook and Microsoft with 1.5 million object instances, 330 thousand images, 80 categories, 91 stuff categories and 5 captions per image. Here, we use filters to mask our object in context and focus on object categories instead of stuff categories to train our model on the cloud. To fine tune the model, we created a dataset by capturing 720p images using raspberry pi camera of daily objects in our list such as, bottle, cell phone, glass, laptop etc.

These images are captured with different lighting conditions and angles for specific model performance. The final dataflow that happens in real time is the live video feed from the raspberry camera which is broken down into individual frames and resized into images of 128p.

D. Model: Mobilenet V2

The Mobilenet v2 model is a type of deep neural network that is designed to be lightweight and efficient for use on mobile and embedded devices. It was developed by Google as an improvement over the original Mobilenet model, with a focus on achieving better accuracy while maintaining a small model size and low computational cost. The Mobilenet v2 model is based on the idea of "inverted residuals," which involves using a series of bottleneck layers with very few parameters followed by a linear layer to increase the model's depth without adding too many parameters as depicted in Fig [3.1]. This approach allows the model to be both lightweight and accurate.

IV. TRANSFER LEARNING

Transfer learning plays a pivotal role in the realization of our project's objectives. In this section, we elucidate the application of transfer learning using TensorFlow Hub, focusing on fine-tuning a pre-trained MobileNet V2 model initially trained on the COCO dataset. This approach allows us to adapt a powerful, general-purpose image recognition model to the specific requirements of our prosthetic arm's object detection and grasp control system.

A. Model Selection

Our choice of the MobileNet V2 model as the base for transfer learning stems from its exceptional efficiency and speed, crucial attributes for our resource-constrained Raspberry Pi-based system. MobileNet V2 has been recognized for its ability to provide remarkable accuracy in object detection while remaining lightweight. This is of paramount importance given the limited computational resources available on a small credit-card-sized computer.

B. Dataset Selection

To fine-tune the model, we require a dataset that aligns with the real-world objects our prosthetic arm will encounter. The "COCO (Common Objects in Context)" dataset stands as an ideal choice due to its extensive collection of approximately 200,000 images spanning around 90 object categories. These images have been meticulously labelled by researchers from prominent organizations such as Microsoft and Facebook. Leveraging this dataset allows us to benefit from a robust foundation of object recognition.

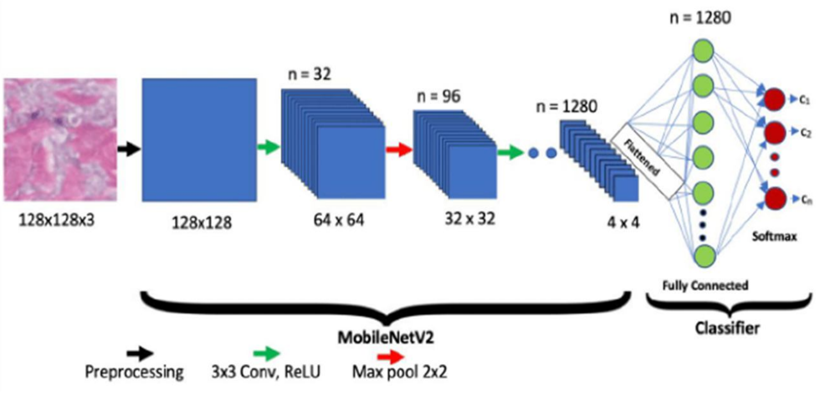

Figure 3.1. Mobilenet Model Architecture

C. Fine-tuning Process

The fine-tuning process begins with the MobileNet V2 model pre-trained on the COCO dataset. We leverage TensorFlow Hub, a repository of pre-trained models and model components, to access the MobileNet V2 model. Subsequently, we fine-tune this model using images captured by the Raspberry Pi camera, adapting it to recognize the specific objects essential for our prosthetic arm's functionality as depicted in the flow diagram given in Fig [4.1]. Fine-tuning involves training the model on our task-specific dataset while retaining the knowledge it gained from its original COCO dataset training. This process ensures that the model generalizes well to our real-world object recognition requirements.

D. Benefits of Transfer Learning

Transfer learning through fine-tuning offers several notable advantages for our project.

- *Efficiency*: By building upon an existing, highly efficient MobileNet V2 model, we harness the benefits of a lightweight architecture suitable for our resource-constrained hardware.
- *Speed*: The fine-tuned model, derived from MobileNet V2, ensures rapid object recognition, enabling real-time grasp control, which is vital for a seamless user experience.
- *Accuracy*: Leveraging a model initially trained on a massive and diverse dataset like COCO provides a strong foundation for recognizing a wide range of daily life objects.
- *Resource Optimization*: Fine-tuning reduces the need for extensive on-device training, conserving both computational resources and time.

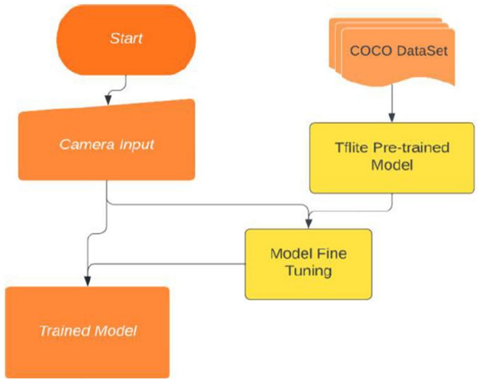

Figure 4.1. Flow Diagram of Model Fine-tuning

V. Experimental Results

In this section, we present the performance metrics, results, analysis, and limitations of our project's experimental phase. Our primary objective was to enable real-time object recognition on a low-cost edge computing device while achieving actionable performance.

A. Performance Metrics

In comparing the performance of MobilenetV1, MobilenetV2, and the modified MobilenetV2 models, several key metrics were observed. MobilenetV1 exhibited an average CPU temperature of 68 °C with passive cooling, which reduced to 54 °C with active cooling. This model achieved a frame rate of 1 frame per second (FPS) and an average inference speed of 300 milliseconds (ms). In contrast, MobilenetV2 demonstrated improved efficiency, with an average CPU temperature of 60 °C (passive cooling) and 44 °C (active cooling), along with an enhanced frame rate of 2.4 FPS and an average inference speed of 380 ms, outperforming MobilenetV1. The modified MobilenetV2 model displayed similar performance characteristics to MobilenetV2, maintaining an average CPU temperature of 62 °C (passive cooling) and 42 °C (active cooling) while consistently achieving a frame rate of 2.4 FPS. Although the average inference speed slightly improved to 320 ms.

B. Results Presentation

Our experiments demonstrated the feasibility of real-time object recognition on edge devices. As per our results as compared in the given table I, Mobilenet V2 provides optimal performance for our task by maintaining the lowest operating temperature while having a competitive frame rate and inference speed.

TABLE I. TABLE OF RESULTS

Model	**Cooling**	**Average CPU temperature (°C)**	**Frame Rate (FPS)**	**Average Inference Speed (ms)**
MobileNet v1	Passive	68	1	300
MobileNet v1	Active	54	1	300
MobileNet v2	Passive	60	2.4	380
MobileNet v2	Active	44	2.4	380
Modified MobileNet v2	Passive	62	2.4	320
Modified MobileNet v2	Active	42	2.4	320

C. Analysis

After careful evaluation, we noticed that Mobilenet V1 demonstrates superior accuracy compared to other models. However, we encountered challenges related to throughput and heat throttling, rendering it impractical for our use case. On the other hand, Mobilenet V2 proved to be a more suitable choice, especially when paired with active cooling mechanisms. It exhibited satisfactory performance and produced actionable results that could be readily utilized. At this point Mobilenet V2 was able to detect all the relevant objects in the frame regardless of how many objects were present. For our specific use case this poses a problem of confusion for the arm grip mechanism. In the pursuit of modification, we then further modified the output layer of our model to always identify and select a single object in a frame. Our model can easily identify an object which is present in the user specified list of six daily life objects as depicted in figure [5.1] and in situations where there are multiple objects of same type present in the frame, our model can choose one with the highest confidence as depicted in figure [5.2].

D. Limitations

- *Confidence Score Dependency*: Our system relies heavily on confidence scores to determine the object of interest and initiate the corresponding grip action. While this approach is effective when dealing with multiple objects of different categories, it may encounter challenges when objects share the same category but possess unique features. In such cases, the system may prioritize the object with the highest confidence score, potentially overlooking distinct object characteristics
- *User-Dependent Decision Making*: In scenarios where, multiple objects belong to different categories (as depicted in Figure 4.7), the system's decision-making process may become complex and potentially confusing. In such situations, we defer the decision to the user wearing the prosthetic arm. The user, through arm positioning, directs the system towards the intended object for grasping or manipulation. Proximity-based grip action is then initiated, grabbing the object brought closest to the prosthetic arm.

Figure 5.1. Object detection with confidence score

Figure 5.2. Object detection for multiple objects with same type

VI. CONCLUSION

In this study, we have successfully developed an efficient on-device object detection system for a prosthetic arm, leveraging state-of-the-art machine learning models. Our primary focus was to achieve real-time object recognition on a low-cost edge device, a critical milestone in enhancing prosthetic functionality. The comparative analysis of MobilenetV1, MobilenetV2, and the modified MobilenetV2 models provided valuable insights into their performance under varying cooling conditions. The results clearly indicate the superiority of MobilenetV2 and its modified version, demonstrating lower average CPU temperatures and significantly improved frame rates and inference speeds. These findings underscore the effectiveness of MobilenetV2 for real-time object recognition in resource-constrained environments.

However, it's important to acknowledge the limitations of our system, particularly in scenarios where objects share similar categories but possess distinctive features. Additionally, user-dependent decision-making may introduce complexities in situations involving multiple objects of different categories. Moving forward, further research and development efforts will be directed towards refining the system's decision-making capabilities and addressing these identified limitations. Additionally, exploring opportunities for real-world deployment and user feedback will be crucial in validating the system's practical utility. Overall, this project represents a significant step forward in advancing the capabilities of prosthetic arms, with the potential to significantly improve the daily lives of individuals with limb loss.

As our project lays a robust foundation for efficient on-device object detection in prosthetic arms, several avenues for future work emerge, aiming to further enhance functionality, usability, and user experience.

- *Hardware Implementation*: Future endeavours will focus on optimizing the hardware components of the prosthetic arm, ensuring seamless integration of the machine learning system. This includes exploring advanced materials and miniaturized components to reduce size and weight while maintaining durability and performance.
- *Efficient Power Management*: To extend battery life and reduce power consumption, research will concentrate on sophisticated power management solutions. Implementing energy-efficient hardware and software techniques will be paramount to ensure prolonged usability without frequent recharging.
- *Effective Cooling Mechanisms*: Enhancing cooling mechanisms will be a priority to maintain optimal operating temperatures, especially during extended use. Innovative cooling techniques, such as passive and active cooling systems, will be explored to prevent overheating and ensure system reliability.
- *Aesthetic Customization*: Recognizing the importance of personalization and aesthetics, future work will include customization options for prosthetic arm users. This will encompass the design and integration of customizable covers and appearances to cater to individual preferences and needs.
- *Connectivity Options*: The integration of versatile connectivity options will be a focal point, enabling seamless communication with external devices and networks. Research will delve into Bluetooth, Wi-Fi, and other wireless technologies to enhance the prosthetic arm's connectivity and functionality.

ACKNOWLEDGMENT

We would like to express our sincere appreciation to Dr. RM Mahalinge Gowda, Principal of PESCE, Mandya, for graciously providing the necessary facilities and resources essential for the successful culmination of this project. Our profound thanks are due to Dr. M L Anitha, Head of the Department of Information Science and Engineering, for her unwavering encouragement, valuable insights, and continuous support throughout the duration of this project.

Special recognition goes to our project guide, T M Geethanjali, Assistant Professor in the Department of Information Science and Engineering, whose guidance and assistance were instrumental in the project's completion. We are also grateful to all the teaching and non-teaching staff whose cooperation and support contributed significantly to the realization of our project goals.

Last but not least, we extend our heartfelt thanks to our parents and friends for their unwavering support and encouragement throughout this endeavor. Each of these contributions has played a vital role in the successful execution of this project, and we acknowledge them with immense appreciation.

REFERENCES

[1] Scott, A. Simon, P. Grochowicz, R. Bibb, and L. Kenney, "Improving prosthetic functionality through sensor fusion and machine learning," Medical Engineering & Physics, vol. 45, pp. 78-84, 2017.

[2] W. Schweitzer, S. M. Posey, and J. A. Strelzow, "The advantages and limitations of body-powered prostheses," Journal of Prosthetics and Orthotics, vol. 21, no. 2, pp. 95-104, 2009.

[3] K. J. Zuo, J. L. Olson, and M. J. Morhart, "Functional outcomes and satisfaction survey of the Bebionic hand prosthesis," Prosthetics and Orthotics International, vol. 40, no. 1, pp. 113-119, 2016.

[4] P. Weiner, A. Galpin, and S. Dubowsky, "Semi-autonomous prosthesis for activities of daily living: preliminary tests with an amputee," Prosthetics and Orthotics International, vol. 39, no. 6, pp. 467-476, 2015.

[5] M. C. Carrozza, M. R. Cutkosky, E. Cavallaro, and M. Zecca, "Control of multifunctional prosthetic hands by processing the electromyographic signal," Critical Reviews in Biomedical Engineering, vol. 34, no. 5, pp. 413-451, 2006.

[6] D. Farina, I. Vujaklija, M. Sartori, T. Kapelner, F. Negro, and N. Jiang, "The neurally controlled bionic hand prosthesis," Frontiers in Neuroscience, vol. 8, p. 33, 2014.
[7] M. A. Puskar and M. J. Howard, "Advances in upper limb prosthetics: a technology review," Medical Devices (Auckland, NZ), vol. 8, pp. 357-366, 2015.
[8] N. Seitz, S. Dosen, and D. Farina, "Hybrid upper limb prostheses: taxonomy and review," Annals of Biomedical Engineering, vol. 44, no. 7, pp. 2078-2090, 2016.
[9] Lake, D. Vescovo, M. Cannizzaro, and N. Donaldson, "The role of 3D printing in prosthetic design and manufacturing," The Journal of Prosthetic Dentistry, vol. 119, no. 4, pp. 545-551, 2018.
[10] E. M. Skidmore, R. O. Smith, and T. A. Peacock, "The impact of upper limb prosthesis use and abandonment on the quality of life and functional outcomes of individuals with congenital and acquired upper limb differences: a systematic review," Disability and Rehabilitation, vol. 39, no. 7, pp. 649-662, 2017.
[11] J. N. Sosa, E. F. Saavedra, and E. A. Finkelstein, "Economic burden of amputation: a systematic review of cost studies," Disability and Health Journal, vol. 8, no. 1, pp. 11-17, 2015.
[12] Castellini et al., "Proceedings of the first workshop on peripheral machine interfaces: going beyond traditional surface electromyography," Frontiers in Neurorobotics, vol. 8, p. 22, 2014.
[13] J. N. Huber, T. Klein, M. Kramer, and P. Blenkhorn, "Virtual Reality in Prosthetic Training and Rehabilitation: A Review of its Effectiveness, Design, and Implementation," Assistive Technology, vol. 30, no. 5, pp. 243-252, 2018.
[14] COCO Dataset: Common Objects in Context (COCO) Dataset, [Online]. Available: https://cocodataset.org/.