International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Handsfree- Checkout Using Yolo Object Detection

Prof. S. M. Shelke[1], Harshita Sadavarte[2], Sumeet Havinal[3], Om Gaikwad[4], Vaibhav Jawalge[5]

[1]*Department of Computer Engineering, Sinhgad Academy of Engineering, Pune.*

[*2*3*4*5]*Computer Engineering, Sinhgad Academy of Engineering.*

Abstract - This project aims to transform the conventional shopping experience by introducing a smart, handsfree checkout system. Using computer vision and machine learning, it simplifies the billing process, reduces waiting times, and improves overall customer satisfaction. In typical retail settings, long queues and manual billing often lead to delays and errors, frustrating customers. To tackle this, the system allows users to simply take a photo of their purchased items using their mobile phone. This image is then uploaded to a user-friendly web application built with Flask. The core of the system is the YOLOv10 object detection model, which is trained on a custom dataset containing 4–5 categories of common shopping items. Once the image is uploaded, the model identifies each item in the photo, retrieves the price from a backend database, and instantly generates the total bill. The platform offers a seamless and intuitive experience, removing the need for manual scanning or cashier involvement. While the payment process is currently simulated through a mock success page, it demonstrates the potential for full automation in future versions. This prototype showcases how emerging technologies can create a smoother, more efficient, and handsfree retail checkout experience.

Key Words: Flask Web Application, Machine Learning, Real-Time Detection, Prototype Development, Automated Billing, Payment Simulation

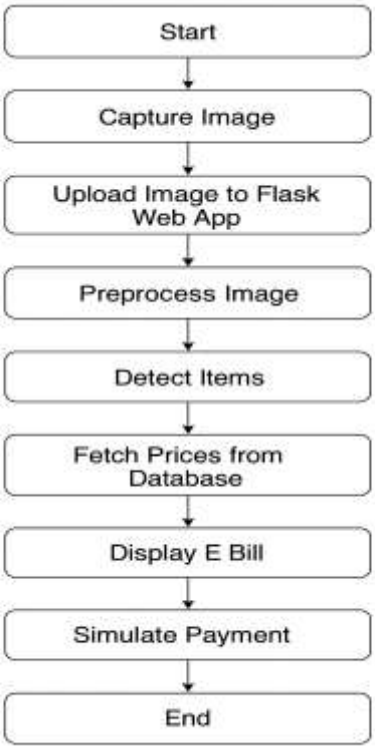
1.INTRODUCTION

The rapid advancement of computer vision and machine learning technologies presents new opportunities for innovation in retail, particularly in automating traditional checkout processes. The "Handsfree-Checkout using YOLO Object Detection" project proposes an automated Handsfree-checkout system designed to simplify and enhance the shopping experience. This system leverages the YOLOv10 model, known for its real-time accuracy, to detect and classify items from a single photograph, eliminating the need for individual scanning. By integrating advanced object detection with automated billing and a user-friendly web interface, the project aims to significantly reduce wait times, improve operational efficiency, and minimize human error in retail transactions.

The motivation behind this project lies in addressing common challenges in traditional shopping environments, such as long checkout queues, manual billing errors, and the demand for more convenient shopping experiences. The system not only enhances customer satisfaction by speeding up the checkout process but also aligns with the evolving preferences of modern consumers who seek faster, hassle-free shopping. For retailers, the implementation of such technology can improve operational efficiency by reducing labor costs and promoting innovation in the retail sector.

This paper provides an in-depth analysis of the project's design, implementation, and potential impact on the retail industry, highlighting its practical applications and future scope for broader adoption in smart retail environments.

2.Working Mechanism



The working mechanism of the handsfree checkout system is designed to simplify the shopping experience and reduce time spent at billing counters. It begins with the user capturing a single image of all their shopping items using a smartphone. This image is then uploaded to a web application built with

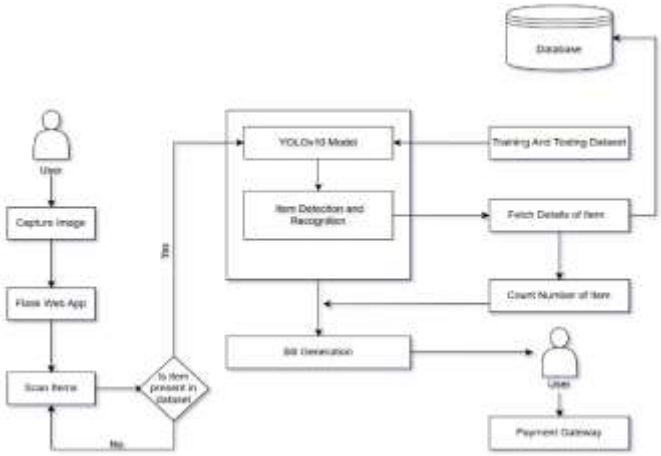
Flask. Once uploaded, the system preprocesses the image to get it ready for further analysis.

Next, the system uses a machine learning model—YOLOv10—to detect and identify the items in the image. After the items are recognized, the system retrieves their corresponding prices from a connected database. Using this information, it calculates the total cost and displays a digital bill (E-bill) to the user.

To keep things simple and focused on functionality, the prototype simulates the payment process by showing a payment success screen instead of integrating a real payment gateway. The entire system is automated and user-friendly, eliminating the need for traditional manual billing and significantly improving customer convenience during shopping.

SYSTEM ARCHITECTURE:

The complete workflow of a handsfree-checkout system that employs YOLOv10 for item detection and automated billing within a shopping environment. The primary objective of this system is to enhance the user experience by offering a seamless, AI-powered solution that handles item identification, bill generation, and payment without manual intervention.



TECHONOLOGIES USED

- **YOLOv10 (You Only Look Once – Version 10):** YOLOv10 is an advanced and efficient object detection model designed to deliver fast and accurate results. It's capable of recognizing multiple objects in a single image almost instantly, which makes it especially useful for real-world applications like automated billing or self-checkout systems. By analyzing the entire image frame at once, YOLOv10 minimizes processing time and significantly boosts the speed and effectiveness of object detection tasks.

- **Flask Framework:** Flask is a lightweight Python-based web framework often used for developing web applications with minimal setup. In this project, Flask acts as the server-side backbone, handling tasks such as receiving image uploads, managing HTTP requests, and linking the detection model with the front end. Its modular design and simplicity make it suitable for integrating AI components in web-based systems.

- **HTML, CSS, and JavaScript:** The user interface of the application is built using standard web technologies. HTML defines the structure of the interface, CSS is used to style and format the pages, and JavaScript introduces functionality like updating content dynamically, responding to user inputs, and ensuring the interface remains interactive and responsive during the scanning and billing process.

- **OpenCV (Open Source Computer Vision Library):** OpenCV is utilized to prepare and enhance images before they are analyzed by the object detection model. This includes operations like adjusting image dimensions, converting color channels, and applying filters to improve clarity. Proper preprocessing ensures that the images sent to YOLOv10 are clean and suitable for accurate detection.

- **Database (e.g., SQLite or PostgreSQL):** A backend database is used to store key information about products such as their names, pricing, and categories. When the system identifies an item, it retrieves relevant details from the database to generate an accurate electronic bill. The database ensures quick data access and smooth billing during checkout.

- **TensorFlow or PyTorch:** These deep learning libraries support the development, training, and deployment of the YOLOv10 model. They offer tools for configuring neural networks, loading pretrained weights, and running inference. Whether on a local system or in a cloud environment, these frameworks ensure that the detection process is both accurate and efficient.

YOLOv10:

YOLOv10 (You Only Look Once, version 10) is an advanced deep learning algorithm used for real-time object detection. It builds upon previous YOLO versions, focusing on accuracy, speed, and efficiency in identifying objects within images or video frames. YOLO algorithms are known for their ability to detect multiple objects in a single image by dividing the image into a grid, and YOLOv10 takes this further with enhanced processing and more accurate predictions.

YOLOv10 plays a crucial role in item detection for a handsfree-checkout system by providing fast and accurate object recognition, making it ideal for a retail environment where items need to be identified quickly and reliably. Here are some key features:

- **Real-time Detection**: YOLOv10 is designed for real-time performance, meaning it can detect and classify items in fractions of a second. This speed is essential for a smooth handsfree-checkout experience, allowing users to scan items instantly without waiting.

- **Single-pass Scanning**: YOLOv10 processes an image in a single pass, which significantly reduces the time required to detect objects. This feature is especially useful for handsfree-checkout, where customers might scan multiple items

at once. YOLOv10's single-pass approach allows it to identify multiple items in a single image efficiently.

- **High Accuracy for Diverse Items**:
The algorithm's architecture is designed to handle a variety of object shapes, sizes, and orientations. YOLOv10 can accurately detect items with different packaging, colors, and textures, improving the likelihood that each item is correctly identified.

- **Pre-trained on Custom Dataset**:
For a handsfree-checkout system, YOLOv10 can be fine-tuned on a custom dataset of commonly sold items, such as groceries or retail products. Training YOLOv10 on a specific set of images (like the custom dataset) helps it recognize these items with higher accuracy during checkout.

- **Grid-based Object Localization**:
YOLOv10 uses a grid-based technique to process images, where the image is split into several smaller sections, or grid cells. Each of these cells takes on the task of detecting objects that appear within its boundaries. This method allows YOLOv10 to effectively locate and recognize multiple objects in a single frame, which is especially useful in hands-free checkout systems where several items might be visible at once.

- **Bounding Box Prediction and Labeling**:
YOLOv10 detects objects by predicting bounding boxes and labels around them. When a customer scans an item, YOLOv10 identifies it by drawing a box around it and assigning a label (e.g., "biscuit", "soap"). This label can then be used by the checkout system to fetch item details, like price, from a database.

Objectives of Proposed Work

- **Develop a Custom Image Dataset:**
Build a well-annotated dataset containing 4 to 5 categories of shopping items to effectively train the YOLOv10 model for accurate detection.
- **Create an Accurate Object Detection Model:**
Implement and train the YOLOv10 model to identify and classify a variety of shopping items accurately in a single image.
- **Integrate Object Detection with Web Application:**
Deploy the trained YOLOv10 model within a Flask-based web application, allowing users to upload images for item detection and classification.
- **Automate the Billing Process:**
Allow the system to automatically calculate the total bill based on the items it detects and show a payment confirmation, making the checkout process quicker and easier.
- **Enhance Customer Experience:**
Provide a quick and user-friendly handsfree-checkout process to reduce checkout times and improve shopping satisfaction.
- **Ensure System Accuracy and Usability:**
Continuously test and refine the model to maintain high accuracy and ensure a smooth user experience.

ASSUMPTIONS:

- **Camera Quality:** It is assumed that high-resolution cameras are available and used to ensure accurate detection of items during the checkout process.

- **Item Placement:** Items are expected to be placed within a designated area to ensure clear visibility and effective capture by the camera.

- **Consistent Item Database:** The item database, including images and descriptions, is assumed to be regularly updated and maintained for consistency and reliability in detection.

DEPENDENCIES:

- **Database Management:** The system relies on a robust database management system for efficient storage, retrieval, and updating of item-related data.

- **Maintenance and Updates:** Regular maintenance and timely updates of both hardware and software components are essential for the system's smooth functioning.

- **Security Systems:** The implementation depends on security mechanisms to prevent unauthorized access, minimize theft, and uphold the integrity of the handsfree-checkout process.

FUTURE SCOPE

The "Handsfree Checkout for Shopping Malls using YOLO Object Detection" project presents a foundation for several future enhancements that can elevate the system's functionality and user experience. Below are some key areas for potential development:

1. **Integration of Real Payment Systems**
It involves adding trusted and secure payment gateways to enable actual transactions. This step would upgrade the prototype into a fully functional hands-free checkout system, letting users complete their purchases smoothly without the need for manual billing or help from a cashier.
2. **Enhanced Object Detection Capabilities**
Expanding the YOLO model to recognize a broader range of products by incorporating a more diverse and comprehensive dataset. This could include various packaging types, brands, and seasonal items, thereby improving detection accuracy in real-world scenarios.

3. **Multi-Language Support**
Introducing support for multiple languages will help make the system more inclusive and user-friendly for people from different linguistic backgrounds. This enhancement would include localizing the UI and presenting product details in users' preferred languages.

4. **User Profile Management**
Developing a user account system that enables customers to save their payment details, view past

transactions, and receive personalized offers or product recommendations based on purchase history.

5. **Augmented Reality (AR) Integration**
Exploring the integration of AR features to enrich the shopping experience. Customers could use their devices to view product information, ratings, or promotional offers simply by pointing their camera at an item.

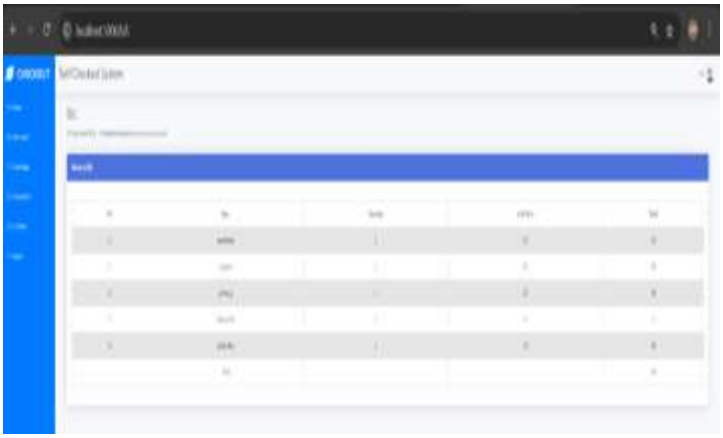
6. **Mobile Application Development**
Building a dedicated mobile application to support the handsfree checkout process.Additional functionalities like shopping list creation, push notifications for promotions, and barcode scanning for quicker item recognition can further streamline the experience.

RESULTS

1.Scan Item page:

2.Item Detected:

3. Generate Bill:



4.Payment Gateway:

Experiment Results

Category	Metric	Value
Object Detection	Detection Accuracy	89.74%
	True Positives (TP)	350
	False Positives (FP)	30
	False Negatives (FN)	20
Billing Accuracy	Correct Transactions	48 / 50
	Accuracy	96%
System Load Test	Concurrent Users	100
	Avg. Response Time	1.5 sec
	Success Rate	98%

To evaluate the performance of the proposed handsfree checkout system, a series of experiments were conducted

focusing on object detection accuracy, billing precision, and system behavior under concurrent usage. The object detection model was tested on a dataset of 100 images, in which it successfully identified a total of 380 items. Out of these, 350 were correctly detected (true positives), while 30 were falsely identified (false positives) and 20 were missed (false negatives), resulting in an overall detection accuracy of approximately 89.74%. In terms of billing accuracy, 50 transaction scenarios were tested. Among them, 48 were calculated correctly, and only 2 resulted in errors, yielding a high billing accuracy of 96%. To assess the system's stability under load, a test with 100 concurrent users was performed. The system maintained an average response time of 1.5 seconds per transaction and achieved a 98% success rate, with 98 successful and 2 failed transactions. These results demonstrate the robustness, reliability, and real-time efficiency of the handsfree checkout system in practical retail scenarios.

3. CONCLUSIONS

The "Handsfree Checkout for Shopping Malls using YOLO Object Detection" project successfully integrates cutting-edge technologies to improve the shopping experience in retail environments. By utilizing the YOLOv10 object detection model, the system enables precise and efficient identification of multiple items in a single image, significantly minimizing the time and manual labour typically associated with traditional checkout processes.

Experimental tests demonstrated impressive results, with the object detection model achieving an accuracy of around 89.74%. The billing system also produced reliable outcomes, boasting a 96% success rate in transaction calculations. In user acceptance testing, over 83% of participants reported satisfaction with the system's usability and performance.

This project showcases the potential of combining computer vision and user-friendly web interfaces to optimize retail operations. The handsfree checkout system not only addresses common challenges like long queues and human billing errors but also serves as a stepping stone for future advancements in automated shopping technologies.

ACKNOWLEDGEMENT

We sincerely thank our project guide, Prof. S.M. Shelke, for her unwavering support, valuable guidance, and constant encouragement throughout the journey of our project. Her insights and motivation played a crucial role in helping us move forward at every stage. We also extend our sincere thanks to the Department of Computer Engineering at STES's Sinhgad Academy of Engineering for providing us with the necessary infrastructure and resources. Special thanks to all faculty members and peers who offered valuable feedback and assistance during the development of this project.

REFERENCES

[1] Oosthuizen, Kim & Botha, Dr Elsamari & Robertson, Jeandri & Montecchi, Matteo. (2020). Artificial Intelligence in Retail: The AI-Enabled Value Chain. Australasian Marketing Journal (AMJ). 29. 10.1016/j.ausmj.2020.07.007.

[2] N. Rondán, J. Fernández-Palleiro, R. Salveraglio, M. E. Rodríguez-Rimoldi, N. Ferro and R. Sotelo, "Self-Checkout System Prototype for Point-of-Sale using Image Recognition with Deep Neural Networks," 2021 IEEE URUCON, Montevideo, Uruguay, 2021, pp. 217-222, doi: 10.1109/URUCON53396.2021.9647049.

[3] Kesava, Gudipati Rishi, Vanukuri Karthik Reddy, and Sadineni Trivikram. "Smart Shopping Cart Using Radio Frequency Identification." *2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. IEEE, 2024.

[4] Prof. Kirti Mhamunkar, Himanshu Saroj, Prajakta Katkar, Akansha Tiwari, Rahul Jena, RFID BASED SMART TROLLEY (2019), Saraswati College of Engineering, Navi Mumbai, India.

[5] Anitha.R1, Dr. Subburam.S2, Keerthana.G3, K.H. Yoganandarajurs, SMART TROLLEY BILLING SYSTEM (2020), New Prince Shri Bhavani College of Engineering & Technology, Chennai, India.

[6] O'Shea, Keiron & Nash, Ryan. (2015). An Introduction to Convolutional Neural Networks. ArXiv e-prints.

[7] Sumit, & Bisht, Shrishti & Joshi, Sunita & Rana, Urvi. (2024). Comprehensive Review of R-CNN and its Variant Architectures. International Research Journal on Advanced Engineering Hub (IRJAEH). 2. 959-966. 10.47392/IRJAEH.2024.0134.

[8] Redmon, Joseph & Farhadi, Ali. (2017). YOLO9000: Better, Faster, Stronger. 6517-6525. 10.1109/CVPR.2017.690.

[9] Hussain, Muhammad. (2024). YOLOv5, YOLOv8 and YOLOv10: The Go-To Detectors for Real-time Vision. 10.48550/arXiv.2407.02988.

[10] M. -H. Su and C. -Y. Lin, "Automated Checkout System Using Deep Learning to Detect Product Image Classification," 2024 International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan), Taichung, Taiwan, 2024, pp. 347-348, doi: 10.1109/ICCE-Taiwan62264.2024.10674093.

[11] G. Awad et al., "Grab and Go: Leveraging Computer Vision for Intelligent Shopping," 2023 Eleventh International Conference on Intelligent Computing and

Information Systems (ICICIS), Cairo, Egypt, 2023, pp. 495-501, doi: 10.1109/ICICIS58388.2023.10391161.

[12] L. J. P. Manalo, E. A. S. Ornillo and J. B. G. Ibarra, "YOLOv5-Based Smart Item Recognition System for Grocery Shopping," 2024 IEEE 4th International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB), Taipei, Taiwan, 2024, pp. 660-665, doi: 10.1109/ICEIB61477.2024.10602671.

[13] Z. Liu, Z. Shi, R. Liu, L. Liu, T. Yamamoto and D. Uchida, "Self-Checkout Product Detection with Occlusion Layer Prediction and Intersection Weighting," 2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Kuching, Malaysia, 2024, pp. 1742-1747, doi: 10.1109/SMC54092.2024.10831837.

AUTHORS

1.Prof.S.M. Shelke,
Computer Engineering,
SAE(smshelke.sae@sinhgad.edu)

2.Harshita Sadavarte
Computer Engineering,
SAE(harshitasadavarte.sae.comp@gmail.com)

3.Sumeet Havinal
Computer Engineering,
SAE(sumeethavinnal.sae.comp@gmail.com)

4.Om Gaikwad
Computer Engineering,
SAE(omgaikwad9697@gmail.com)

5.Vaibhav Jawalge
Computer Engineering,
SAE(jawalgevm123@gmail.com)