VIVA-TECH IJRI
ANNUAL RESEARCH JOURNAL

VIVA-TECH INTERNATIONAL JOURNAL FOR RESEARCH AND INNOVATION
ANNUAL RESEARCH JOURNAL
ISSN(ONLINE): 2581-7280

SURVEY ON AUTOMATED CHECKOUT SYSTEM

Darshan Dhamode[1], Danish Shaikh[2], Ranjit Narlya[3] , Prof. Kirtida Naik[4]

[1](Department Computer Engineering, Mumbai University, India)
[2](Department Computer Engineering, Mumbai University, India)
[3](Department Computer Engineering, Mumbai University, India)
4 (Professor, Department Computer Engineering, Mumbai University, India)

Abstract: *This research focuses on an advanced automated checkout system using the YOLOv8 model for real-time object detection. By capturing product images at checkout, the system identifies items without traditional barcode scanning. Trained on a custom dataset, YOLOv8 detects and classifies products, distinguishing between similar packaging variants. It retrieves product details from a database, updates the cart, and calculates the total for a seamless checkout experience. This solution improves speed, reduces errors, and eliminates the need for costly scanning hardware, making it adaptable for various retail setups. By leveraging image-based detection, the system offers flexibility and high accuracy, even during peak store traffic. Additionally, it streamlines inventory tracking, enhances operational efficiency, and improves overall customer satisfaction. Integrating machine learning, computer vision, and object detection, this research focuses on revolutionizing the retail experience.*

Keywords - *YOLOv8 Model, Image-Based Product Detection, Vision-Based Automation*

I. INTRODUCTION

This research focuses on an automated checkout system leveraging advanced computer vision and machine learning techniques to streamline retail processes. Using the YOLOv8 model for real-time object detection, it identifies products from images instead of relying on barcodes. The system accurately classifies items, even with similar packaging, by recognizing attributes such as size or weight. Designed for high-traffic environments, it enhances speed, reduces human error, and eliminates the need for costly hardware. By capturing images at the checkout counter, the YOLOv8 model processes them in real time to detect and classify products. The model is trained on a custom dataset featuring various retail items, enabling it to recognize and differentiate products with high precision. Once items are identified, the system calculates the total price, providing a seamless and efficient experience for customers and staff alike. This image-based detection approach offers flexibility by eliminating barcodes, which can be damaged or obscured, and supports a broader range of product packaging. The system is scalable, easily accommodating new products without requiring hardware modifications, making it ideal for diverse retail setups.

II. REVIEW OF LITERATURE SURVEY

A survey was done on the existing literature and products to find out their shortcomings and research gaps in their systems. This survey consisted of 12 literature papers where the most relevant ones are listed below.

Svetlana Illarionova Sergey Nesteruk Nadezhda Mikhailova, Andrey Somov, Ilya Zherebzov, Claire Traweek and Andivanoseledets [1] introduces a novel approach for smart checkout systems, allowing the recognition of new product classes without requiring human-annotated data. The paper proposes "Pseudo Augment," a data augmentation method that generates pseudo-labels to expand datasets automatically. This technique enables machine learning models to generalize and recognize previously unseen items, improving scalability and reducing the need for costly manual labeling. The research showcases its application in retail environments, enhancing checkout systems by facilitating quick adaptation to new products.

Tzu-Wei Yu1, Muhammad Atif Sarwar and Sheng-Hsien Cheng [2] explores a novel approach to understanding and analyzing activities in dynamic environments by leveraging spatiotemporal semantics. The proposed iCounter framework focuses on the segmentation of foreground objects, separating them from background elements to detect and track objects or people over time. By utilizing spatiotemporal data, the system captures both spatial and temporal information about object movements and interactions, allowing for a more detailed understanding of activity patterns. This method enhances the accuracy of counting and tracking in scenarios such as retail environments, surveillance, and crowd monitoring. The framework aims to improve event detection and object tracking by integrating object segmentation with semantic analysis, making it valuable for applications requiring precise activity understanding and monitoring over time.

Mohammed alghamdi Mohammed maray ,Hanan abdullah mengash and Mohammed aljebreen [3] presents a comprehensive study on enhancing consumer product recognition in retail environments through the integration of the Aquila Optimization Algorithm with deep learning techniques. The research addresses the challenges of accurately identifying products on store shelves, which is crucial for inventory management and customer experience. By leveraging the Aquila Optimization Algorithm, the study optimizes the performance of deep learning models, improving their ability to recognize diverse product categories in real-time. The approach emphasizes the importance of high accuracy and efficiency in product recognition systems, showcasing how advanced algorithms can significantly reduce computational costs and enhance detection rates. Through extensive experiments, the paper illustrates the effectiveness of this optimized model in various retail scenarios, ultimately demonstrating its potential to empower retailers by providing actionable insights into consumer behavior and inventory management. The findings contribute to the growing field of computer vision in retail, paving the way for smarter, more responsive retail operations.

Bing-Fei Wu, Wan-Ju Tseng, Yung-Shin Chen, Shih-Jhe Yao and Po-Ju Chang [4] introduces a cutting-edge self-checkout solution designed to enhance the retail shopping experience by integrating advanced technologies such as computer vision, machine learning, and user-friendly interfaces. The system aims to streamline the checkout process, reducing wait times and improving customer satisfaction. By employing sophisticated object recognition algorithms, the intelligent self-checkout system can accurately identify and process a wide variety of

products, even those with similar packaging, minimizing errors during checkout. Additionally, the system leverages real-time data analytics to monitor shopping patterns and inventory levels, enabling retailers to optimize stock management and improve operational efficiency. The paper discusses the architecture and implementation of the system, highlighting its adaptability to various retail environments, from grocery stores to convenience shops. Through comprehensive testing, the research demonstrates significant improvements in transaction speed and accuracy, suggesting that such intelligent systems can transform the retail landscape by making the checkout experience seamless and more enjoyable for customers.

M. F. M. Busu, I. Ismail, M. F. Saaid, S. M. Norzeli [5] presents an innovative solution for automating the checkout process in retail environments through the use of RFID technology. The paper discusses the challenges associated with traditional checkout systems, such as long wait times and human errors, and proposes an RFID-based auto-checkout system that enhances efficiency and accuracy. By embedding RFID tags in products, the system allows for quick and automatic identification of items as customers place them in their carts. This seamless tracking facilitates real-time inventory management and significantly reduces the need for manual scanning at checkout points. The research details the system's architecture, which integrates RFID readers, a centralized database, and a user-friendly interface, allowing customers to check out with minimal effort.

M.A. Besbes, H. Hamam, [6] explores the development and implementation of an advanced checkout system utilizing Radio Frequency Identification (RFID) technology to enhance retail operations. The paper addresses the limitations of traditional checkout methods, such as lengthy wait times and inaccuracies during product scanning. By integrating RFID technology, the proposed system allows for automatic identification and tracking of products, enabling a seamless and efficient checkout experience.

Sangwon Hwang, Jisun Lee and Seungwoo Kang [7] presents a framework that combines text detection with computer vision algorithms, allowing for real-time recognition and tracking of products in a dynamic environment. This integration facilitates a seamless AR experience, where users can view additional product details, promotions, or even user reviews overlaid on the physical items. The authors conduct a series of experiments to evaluate the performance of the system in various retail scenarios, demonstrating its effectiveness in accurately recognizing and tracking products based on textual information.

Prabu Selvam and Muhammad Faheem [8] introduces an innovative neural network architecture specifically designed for grocery product recognition, eliminating the need for batch normalization. The research addresses the limitations of traditional convolutional neural networks (CNNs) that often rely on batch normalization to stabilize training and improve performance. Instead, this study proposes a feature flow neural network that emphasizes rigorous feature extraction and flow, allowing for effective learning without the computational overhead of batch normalization.

Zhang Taoning and Chen Enqing [9] presents a novel approach to product recognition by integrating Histogram of Oriented Gradients (HOG) features with a Bag of Words (BoW) model. This research addresses the challenges of accurately identifying products in diverse retail environments, where variations in lighting, angles, and packaging can hinder recognition efforts. The HOG feature extraction method is employed to capture essential shape and edge information from product images, providing a robust descriptor that enhances the algorithm’s ability to distinguish between different items. The paper outlines the process of combining HOG

features with the BoW model, which quantizes the feature space into discrete visual words, allowing for efficient representation and comparison of product images.

Yuchen Wei, Shuxiang Xu, Son Tran, and Byeong Kang [10] explores the use of Generative Adversarial Networks (GANs) as a powerful technique for augmenting datasets in the domain of grocery product image recognition. The study addresses the common challenge of limited labeled data, which can hinder the performance of deep learning models in accurately recognizing products in retail environments. By employing GANs, the research aims to generate synthetic images that closely resemble real product images.

Hotaka Niwa, Koichi Nagata, Masaya Ohta and Katsumi Yamashita [11] presents an innovative approach to enhancing online shopping experiences through the integration of mixed-reality technologies and advanced product identification methods. This research addresses the limitations of traditional e-commerce platforms by enabling users to interact with products in a more immersive and engaging manner. The proposed system leverages computer vision techniques to accurately identify and recognize products in real-time, allowing consumers to visualize items in their own environment through augmented reality (AR).

Sai Neeraj Kanuri, Samarth P Navali and S Rahul Ranganat [12] explores an advanced approach to product recognition and labeling by implementing a multi neural network architecture. This research addresses the challenges of accurately identifying and categorizing a wide range of products in retail environments, where variations in packaging, design, and lighting can complicate recognition tasks. The proposed model leverages the strengths of multiple neural networks, each trained on different subsets of product data, to improve overall recognition accuracy and robustness. The paper outlines the architecture of the multi-neural network system, detailing how each network specializes in identifying specific categories or features of products.

III. ANALYSIS

Table 1: Analysis Table

Title of Paper	Technology	Advantages	Disadvantages
Spatiotemporal Activity Semantics Understanding Based on Foreground Object Segmentation: iCounter Scenario (2024)	1. AI 2.Machine Learning	1. Combination of remote sensing and machine learning techniques.	1. May require significant computational resources.
Empowering Retail Through Advanced Consumer Product Recognition Using Aquila Optimization Algorithm with Deep Learning (2024)	1. Artificial Intelligence 2. Machine Learning	1. Combination of AI and machine learning techniques	1. High computational costs associated with deep learning models and the need for extensive hyperparameter tuning.

Automated Checkout System Using Deep Learning to Detect Product Image Classification (2024)	1. Deep Learning	1. Accuracy may decrease when dealing with an imbalanced dataset, as the model may favor more represented classes.	1. Systems need to continuously update and retrain models to accommodate new products entering the inventory.
PseudoAugment: Enabling Smart Checkout Adoption for New Classes Without Human Annotation (2023)	1. Deep Learning 2. Machine Learning	1. Improves performance of smart checkout systems	1. Requires top view images only. May not be necessary if sufficient training data is available.
Batch Normalization Free Rigorous Feature Flow Neural Network for Grocery Product Recognition (2023).	1. Embedded IoT Devices	1. Enhances the accuracy of grocery product recognition	1. Developing and fine tuning the network architecture can be challenging.
Enhancing Retail Checkout through Video Inpainting, YOLOv8 Detection, and DeepSort Tracking (2023)	1. Video Inpainting 2. Real-time Video Processing	1. DeepSort ensures smooth tracking of multiple products, even in crowded or fast-moving environments	1. Real-time video inpainting, detection, and tracking require significant computational power, which may not be feasible for all retail setups.
YOLOv8 for Product Brand Recognition as Inter-Class Similarities	1. YOLOv8	1. Fine-tuned YOLOv8 excels at distinguishing products with overlapping visual features, improving inter-class recognition.	1. Performance heavily relies on the quality and diversity of the training dataset, requiring significant effort to annotate and prepare.
Enabling Product Recognition and Tracking Based on Text Detection for Mobile Augmented Reality (2022).	1. OCR	1. High Accuracy in Text Detection.	1. Significant investment required for developing and maintaining AR and OCR technologies.
DeepACO: A Robust Deep Learning-based Automatic Checkout System (2022).	1. Deep Learning	1. DeepACO is likely designed to handle complex retail environments, including variations	1. Struggles to differentiate between products with minor visual differences, such as size or flavor variants.

		in lighting, occlusion, and diverse product arrangements	
Self-supervised Contrastive learning for Zero-shot automatic retail checkout (2022)	1. Self-supervised Learning 2. CNN	1.Self-supervised learning eliminates the reliance on costly and time-consuming manual annotation.	1. Real-world variances like lighting, occlusion, and camera quality may impact accuracy, especially in zero-shot scenarios.
Product Recognition Algorithm Based on HOG and Bag of Words Model (2021).	1. HOG 2. BOW	1. HOG provides detailed edge and gradient information, useful for distinguishing products.	1. HOG and BoW may not capture complex patterns as effectively as deep learning models.
Data Augmentation with Generative Adversarial Networks for Grocery Product Image Recognition (2021).	1. CNN	1. Augmented data can help improve the accuracy and robustness of image recognition models.	1. Training GANs can be complex and computationally intensive.
Improving Deep Learning-based Automatic Checkout System Using Image Enhancement Techniques (2021).	1. Deep Learning Framework	1. Enhancing image quality ensures better feature extraction, leading to more accurate classification and detection.	1. Over-reliance on image enhancement can lead to models that perform poorly on raw, unprocessed images.
Multi neural network model for product recognition and labelling (2021).	1. Internet of Things (IoT) Framework	1. Combining multiple neural networks can enhance the accuracy of product recognition and labeling	1. Increases the complexity of model architecture and training process.
Self-checkout Intelligent Recognition System for Fruit Supermarkets	1. Machine Learning 2. CNN	1. Speeds up the checkout process,	1. Difficulty distinguishing between similar-looking fruits

(2019).		reducing customer wait times.	

IV. CONCLUSION

The development of the automated checkout system represents a significant advancement in retail technology, addressing the limitations of traditional checkout processes and enhancing the overall shopping experience for customers. By integrating cutting-edge technologies such as computer vision, machine learning, and real-time data management, the system aims to streamline product recognition, reduce checkout times, and minimize human error. Through a comprehensive implementation plan that encompasses requirement analysis, system design, development, integration, and user testing, this project lays the foundation for a robust and user-friendly solution. The anticipated outcomes include improved operational efficiency for supermarkets, enhanced customer satisfaction through quicker transactions, and a more interactive shopping experience. Additionally, the system provides scalability, making it adaptable to different store sizes and formats, from small retail outlets to large supermarkets. It also offers real-time inventory tracking and management, helping retailers optimize stock levels and reduce waste. By eliminating the need for traditional barcode scanning, the solution ensures a smoother and more flexible checkout process, where product identification can be achieved even when labels are damaged or obscured.

Acknowledgements

We would like to express our sincere gratitude to the organization for their invaluable support and contribution throughout this project.

REFERENCES

[1] T. -W. Yu, M. A. Sarwar, Y. -A. Daraghmi, S. -H. Cheng, T. -U. İk and Y. -L. Li, "Spatiotemporal Activity Semantics Understanding Based on Foreground Object Segmentation: iCounter Scenario," in *IEEE Access*, vol. 10, pp. 57748-57758, 2024.

[2] M. Alghamdi, H. A. Mengash, M. Aljebreen, M. Maray, A. A. Darem and A. S. Salama, "Empowering Retail Through Advanced Consumer Product Recognition Using Aquila Optimization Algorithm With Deep Learning," in *IEEE Access*, vol. 12, pp. 71055-71065, 2024

[3] 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

[4] S. Nesteruk *et al*., "PseudoAugment: Enabling Smart Checkout Adoption for New Classes Without Human Annotation," in *IEEE Access*, vol. 11, pp. 76869-76882, 2023

[5] P. Selvam *et al*., "Batch Normalization Free Rigorous Feature Flow Neural Network for Grocery Product Recognition," in *IEEE Access*, vol. 12, pp. 68364-68381, 2024

[6] S. Hwang, J. Lee and S. Kang, "Enabling Product Recognition and Tracking Based on Text Detection for Mobile Augmented Reality," in *IEEE Access*, vol. 10, pp. 98769-98782, 2022.

[7] L. H. Pham *et al*., "DeepACO: A Robust Deep Learning-based Automatic Checkout System," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, New Orleans, LA, USA, 2022, pp. 3106-3113

[8] Z. Taoning and C. Enqing, "Product Recognition Algorithm Based on HOG and Bag of Words Model," *2019 8th International Symposium on Next Generation Electronics (ISNE)*, Zhengzhou, China, 2019, pp. 1-3

[9] Y. Wei, S. Xu, S. Tran and B. Kang, "Data Augmentation with Generative Adversarial Networks for Grocery Product Image Recognition," *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Shenzhen, China, 2020

[10] L. H. Pham *et al*., "Improving Deep Learning-based Automatic Checkout System Using Image Enhancement Techniques," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Vancouver, BC, Canada, 2023

[11] S. N. Kanuri, S. P. Navali, S. R. Ranganath and N. V. Pujari, "Multi neural network model for product recognition and labelling," *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, India, 2018.

[12] J. Abyasa, M. P. Kenardi, J. Audrey, J. J. Jovanka, C. Justino and R. Rahmania, "YOLOv8 for Product Brand Recognition as Inter-Class Similarities," *2023 3rd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, Yogyakarta, Indonesia, 2023

[13] A. Vats and D. C. Anastasiu, "Enhancing Retail Checkout through Video Inpainting, YOLOv8 Detection, and DeepSort Tracking," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Vancouver, BC, Canada, 2023

[14] A. Nath, "Self-supervised Contrastive learning for Zero-shot automatic retail checkout," *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Antalya, Turkey, 2022, pp. 1-4.

[15] Y. Zhou, X. Lv, H. Wang and X. Zhou, "Self-checkout Intelligent Recognition System for Fruit Supermarkets," *2024 2nd International Conference on Signal Processing and Intelligent Computing (SPIC)*, Guangzhou, China, 2024.