Produce Sync

Master of Applied Computing

University of Windsor

Advanced Database Topics

Author:

Deon Victor Lobo

1. ABSTRACT

The "Produce Sync" initiative sets up an approach to reduce agricultural waste through the establishment of an online marketplace that connects farmers and consumers in real time. With the help of MongoDB, Spring Boot, and React, the platform guarantees quick transactions and uses innovative computer vision technology to determine how fresh produce is, allowing for the sale of ripe goods. Unique to this project are dynamic pricing strategies that account for product ripeness and shelf life as well as geospatial queries that allow for accurate product localization. To effectively match buyers and sellers, our methodology integrates computer vision with deep learning models such as ResNet for crop grading, as well as MongoDB's geospatial capabilities. The findings point to a more efficient agricultural supply chain resulting in less food waste, and higher farmer profits. "Produce Sync" illustrates the capacity of technology to tackle formidable challenges in the field of sustainable agriculture.

Index Terms: Sustainable Agriculture, Food Waste Management, Computer Vision in Agriculture, Geospatial Queries, Dynamic Pricing Strategies, Machine Learning in Agricultural Applications, Online Marketplaces for Agriculture, Crop Grading using Deep Learning, Agricultural Supply Chain Efficiency, Environmental Impact of Food Production

2. INTRODUCTION

"Produce Sync" aims at using technology to help farmers reduce food waste, which is a major problem in today's world where food sustainability is a major concern.

Due to issues like excessive production, logistics inconsistencies, and the absence of a real-time platform that links farmers directly to consumers, a considerable amount of produce never makes it to the consumer market, even though agricultural practices have improved. After the US and China, wasted food emits 3.3 gigatons of carbon, making it the third-largest greenhouse gas emitter. Based on producer’s prices (excluding fisheries and seafood), the value of agricultural product waste in the food industry amounts to USD 750 billion. Statistical data released by the Food and Agricultural Organization (FAO) reveal that approximately 1.3 billion tons of food, or one-third of the food produced for human consumption, is either lost or wasted as of 2023.[1] This leads to severe financial losses for farmers and contributes to environmental degradation. This waste contributes to unnecessary consumption of water, land, and energy, further straining our planet's resources. The lack of a system that adjusts prices in response to changes in supply and demand and the freshness of products makes the problem even worse, as it causes producers to lose money and leads to wasted food.

"Produce Sync" makes use of MongoDB for data storage and administration, Spring Boot for backend services, and React for an interactive front end. Buyers and local farmers are connected through the integration of the OpenCage API, which allows for precise geospatial queries. It uses Convolutional Neural Networks (CNNs) to evaluate when produce is ripe which helps in dynamic pricing.

"Produce Sync" uses technologies to make the future of farming more sustainable and efficient. This project aims to connect farmers and consumers by using computer vision, geospatial technology, and dynamic data management. This will make sure that the harvest is used wisely. The project aims to not only cut down on waste, but also improve the farming ecosystem by making every part of the food supply chain more efficient and environmentally friendly.

We are moving toward a system that values real-time data, transparency, and connectivity. The project gives farmers the information and tools they need to make decisions about selling their crops by putting them in direct contact with people who buy their food. This could change the agricultural landscape for the better by making it fairer and more balanced. Convolutional Neural Networks (CNNs) are used to find out when fruits and vegetables are ready to be eaten and emphasis is placed on the produce which is ripe and needs to be sold as early as possible to prevent the wastage of food.
"Produce Sync" comes up with new ways to deal with food waste and inefficiencies in agriculture by using technology to make farming more sustainable and effective. This project aims to make the food system more sustainable and fairer by putting farmers in direct contact with customers and using advanced analytics to get the most out of crop sales.

3. LITERATURE / BACKGROUND STUDY

The use of computer vision for automated tomato grading stands out as a significant advancement in agricultural innovation. Research indicates that this technology greatly speeds up the grading process and improves accuracy, resulting in a success rate of 96.47% in differentiating between high-quality and mature tomatoes.[2] This represents a shift from conventional manual grading techniques, offering potential decreases in both time and human error. Though these improvements have been made, two main drawbacks have been found: the system's limited capacity, which is limited to three hundred fruits per hour, and its inability to process images that contain specular reflections. These challenges highlight the need for additional technological advancements to accommodate more diverse physical conditions and larger-scale operations.

Along with these concerns about technology, the need to keep food sustainable is becoming more and more important on a global level. According to the Food and Agricultural Organization of the United Nations, it is projected that the global population will reach around 9.1 billion by the year 2050.[3] Considering this, the organization emphasizes the significant importance of contemporary technologies such as artificial intelligence (AI) in addressing the increasing need for food.[4] By incorporating computer vision into the process of monitoring post-harvest products, it ensures that only high-quality items are delivered to consumers, thus reducing food waste. The effectiveness of computer vision in tasks such as object detection, recognition, and tracking has been significantly improved by the progress made in machine learning and deep learning.[5] These advancements are critical in preserving food quality standards and mitigating the significant problem of food losses caused by the post-harvest overripening of fruits and vegetables.[6]

By utilizing advanced computer vision and machine learning technologies, such as Convolutional Neural Networks (CNNs), our project, "Produce Sync," directly tackles these identified challenges. With the introduction of improved image processing techniques that can efficiently handle specular reflections, we hope to expand the capabilities of grading systems beyond their present limitations. To do this, deep learning architectures like ResNet are used. ResNet has been very successful at solving the vanishing gradient problem and making image recognition more accurate.[7] Assuring the correct grading of agricultural products and making the agricultural ecosystem work better depends on this.

In addition, our project incorporates cutting-edge price-setting methodologies, including segmented and peak-load pricing, in recognition of the criticality of dynamic pricing strategies in the agricultural industry. This facilitates the implementation of pricing strategies that are better suited to market conditions, customer behavior, and crop quality. Our objective is to determine the most precise CNN model for our purpose through an exhaustive comparison of multiple models. This will enable us to price and sell products in accordance with their ripeness and shelf life, thereby reducing food waste and advancing sustainability.

To make self-sufficient regions, we also used MongoDB's Geospatial queries to connect buyers and sellers directly.[8] This approach serves the dual purpose of supporting local businesses and mitigating the greenhouse gas emissions linked to long-distance transportation.

In summary, "Produce Sync" classifies products into discrete quality tiers by integrating cutting-edge technologies like computer vision, deep learning, and dynamic pricing strategies. Our project takes a step ahead in reducing food waste and making the agricultural supply chain more sustainable. It does this by tackling problems that have been pointed out in the literature, like making grading systems more powerful and improved ways to deal with specular reflections. "Produce Sync" wants to promote a direct link between buyers and sellers by carefully comparing CNN models and using geospatial queries. This will help make the future agricultural ecosystem more efficient and long-lasting.

4. PROPOSED MODEL/ METHODOLOGY

The proposed model aims to empower small-scale farmers by leveraging the capabilities of machine learning and computer vision. This innovative application provides farmers with a dynamic platform to adjust produce prices based on various characteristics, including shelf life, thereby reducing losses from unsold products. Our system facilitates direct connections between farmers and nearby buyers, promoting local transactions and decreasing the reliance on long-distance transportation as shown in Figure 1 and Figure 2. The self-sustaining model operates on the principles of supply and demand, fostering a more efficient agricultural ecosystem. To optimize real-time performance, we will conduct a thorough comparison of different CNN architectures, such as ResNet50, ResNet152, AlexNet, GoogLeNet, and, seeking the most efficient solutions for seamless integration into our application.

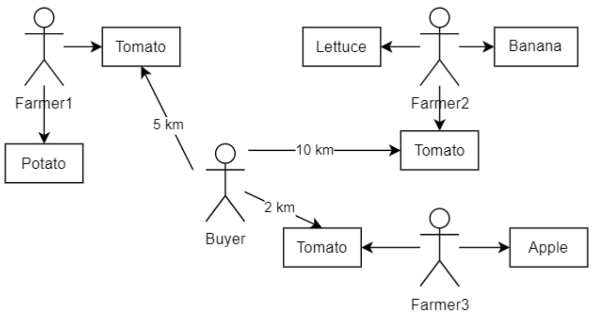

Figure 1. Recommending tomatoes to a buyer within 10km radius

Figure 2. Flowchart of finding all sellers within 10 km radius

Subsequently, we shall implement a strategic segmented pricing system, classifying fruits into distinct ripeness categories: unripe, semi-ripe, ripe, and overripe as shown in Figure 3. This type of pricing model aims to optimize market dynamics by assigning varying values to fruits based on their ripeness. For instance, semi-ripe fruits may command a higher price due to an extended shelf life, while ripe fruits could be priced lower, ensuring quicker turnover and reduced wastage of overripe crops. This type of segmented pricing will facilitate faster sales and reduce the wastage of overripe crops.

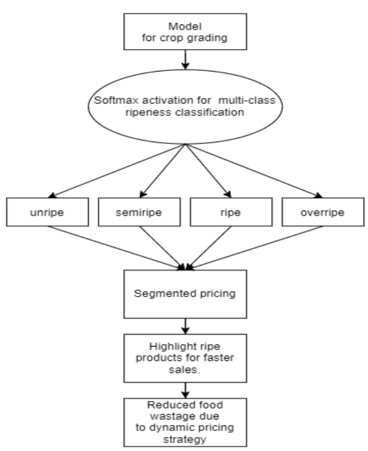

Figure 3. Optimized pricing strategy based on crop grading.

OpenCage will be integrated to obtain accurate coordinates for both buyers and sellers. This external geocoding service will play a pivotal role in translating addresses into precise geographic coordinates. For buyers, the system will identify nearby sellers using MongoDB's geospatial queries, ensuring a streamlined and efficient process. We will utilize MongoDB's aggregation queries geospatial capabilities to facilitate a location-based search for sellers. The $geoNear stage utilizes spherical geometry to locate sellers near a specified point defined by latitude and longitude. Subsequently, it extracts the products matching the buyer's criteria and returns the resulting list of products.

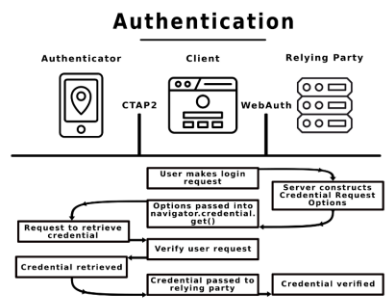

Figure 4. Password less Authentication Flow.

In our implementation, we are adopting a password less authentication [9-10] solution that leverages cryptographic techniques for user access as shown in Figure 4. During the user registration process, a unique key pair, consisting of a public key and a private key, is generated. The private key is securely stored on a designated hardware device, such as a security key or a biometric scanner. When users initiate a login attempt, the server generates a challenge. The user's hardware device utilizes the private key to produce a cryptographic response. Verification occurs on the server side, where the response is authenticated using the stored public key. Successful validation results in user access. This approach not only bolsters security by mitigating password-related vulnerabilities but also enhances user experience through a streamlined and secure authentication process. The system utilizes cryptographic key pairs for each user, enhancing security by eliminating common attack vectors associated with traditional password-based systems.

5. RESULTS

In our cutting-edge React and Spring Boot application, we've created a robust ecommerce platform designed to bridge sellers with local buyers seamlessly. By integrating OpenCage, we've simplified location tracking, effortlessly fetching latitude and longitude data from buyer addresses. Leveraging MongoDB's geospatial queries, our platform efficiently localizes produce, connecting users with nearby farmers, and cutting logistical costs while promoting the consumption of locally sourced goods. The sellers are empowered by our CNN model, which provides intelligent pricing recommendations based on product maturity levels. This ensures swift sales through strategic pricing adjustments, enabling sellers to optimize their pricing strategy for maximum profitability and market competitiveness.

Our implementation leverages password less authentication as shown in Figure 5 and Figure 6 through Spring Security, utilizing JWT tokens for user validation. These tokens, categorized into SELLER and BUYER types, are securely stored within cookies. Unlike traditional methods, our approach eliminates the need for users to input passwords, instead requiring only their email addresses. Upon login attempts, we validate the token stored in the cookies against the corresponding user. If validation fails or if no token exists, we seamlessly generate a new token. This freshly minted token is then dispatched to the user via email for further authentication. With a lifespan of 10 days, our JWT tokens ensure both security and convenience for users, enhancing the overall authentication process.

```
{
  "role": "SELLER",
  "location": {
    "$near": {
      "$geometry": {
        "type": "Point",
        "coordinates": [longitude, latitude]
      },
      "$minDistance": minDistance,
      "$maxDistance": maxDistance
    }
  }
}
```

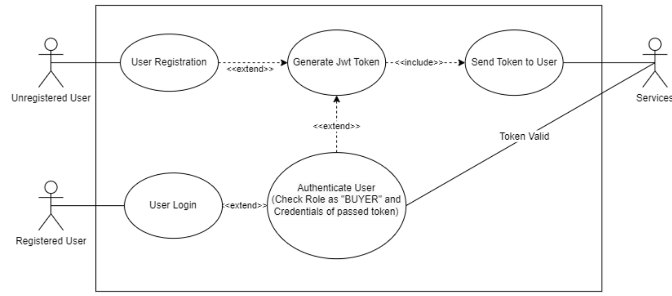

Figure 5. Password less Authentication using spring security.

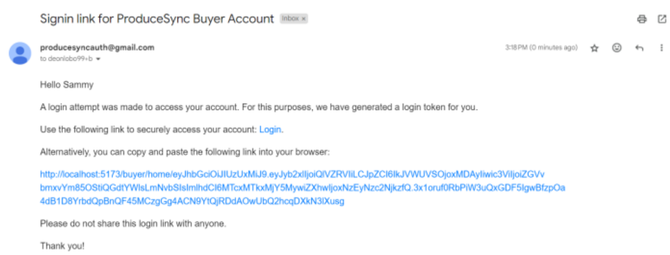

Figure 6. Login email sent to user with token.

This MongoDB query employs the $near operator to precisely locate addresses associated with a specified role within a designated proximity of a geographical point, determined by latitude and longitude coordinates. These coordinates are dynamically obtained from the buyer's address through the OpenCage API, ensuring real-time accuracy in pinpointing nearby sellers. By integrating $minDistance and $maxDistance parameters, the search is further refined to include addresses falling within a specified distance range, in this case, within a 20km radius from the provided coordinates as shown in Figure 7. This approach enables efficient retrieval of sellers located within the desired vicinity, streamlining the process of connecting buyers with nearby merchants.

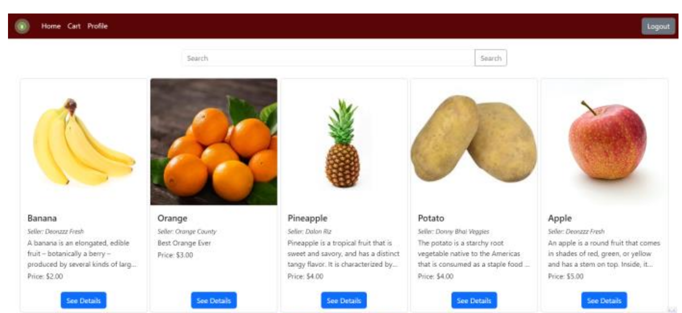

Figure 7. Seller products within 20km radius of the buyer

Once the sellers within the 20km radius of the buyer are extracted, the system seamlessly filters and presents products exclusively from these nearby sellers. This strategic implementation, depicted in Figure 7, enhances user experience by showcasing only products available within the buyer's vicinity, fostering localized transactions, and facilitating quicker access to desired goods. Additionally, a user-friendly search feature empowers buyers to explore specific products of interest, ensuring a tailored browsing experience while promoting efficient and convenient shopping within the defined geographical scope.

The implemented order management system as shown in Figure 8 encompasses a comprehensive suite of features, including authentication, product management, order management, inventory management, and user profile management. Notably, it offers a seamless user experience by facilitating the addition of products from the same seller to the cart, while also intelligently prompting buyers to consider overriding the existing cart contents with products from a different seller. The product page is used to show the details of the available item which is seen in Figure 9. This thoughtful approach ensures ease of logistics and ensures convenience for users while maintaining the integrity of their shopping experience.

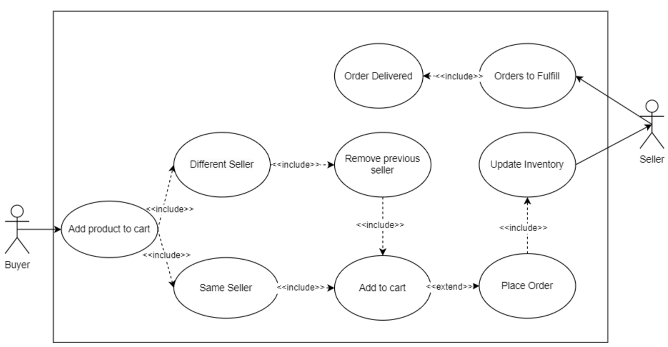

Figure 8. Order Management diagram.

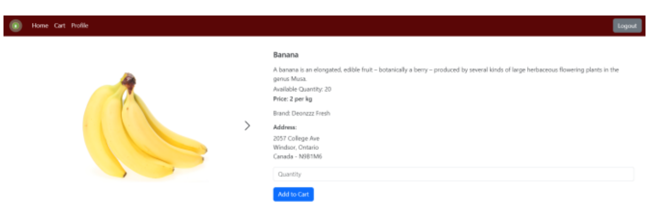

Figure 9. Product page

The cart page provides users with a comprehensive view of all products currently in their cart, enabling them to manage their selections by easily adding or removing items as needed. Each item in the cart retains an "OPEN" status until the user finalizes their purchase. Upon clicking the "confirm order" button, the status of the entire order transitions to "CONFIRMED" indicating that the order has been successfully placed. On confirmation the available quantity of the respective products is automatically adjusted, ensuring the maintenance of an accurate and up-to-date inventory. This proactive measure helps prevent overselling and ensures that customers are only presented with products that are genuinely in stock, thus enhancing the reliability and efficiency of the ordering process. This seamless process ensures clarity and transparency for users, allowing them to confidently proceed with their purchase while maintaining visibility over their shopping decisions.

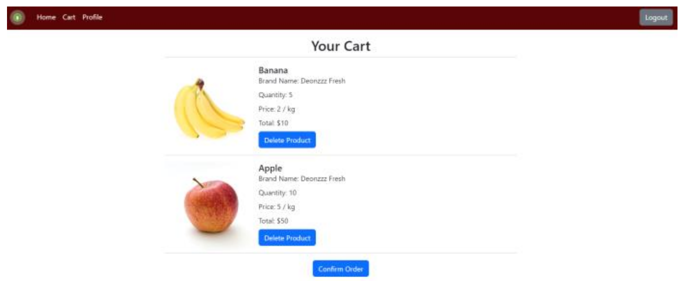

Figure 10. Buyers products in the cart

Within the seller panel, a dedicated section titled "Confirmed Orders" offers sellers a comprehensive overview of all orders awaiting fulfillment as shown in Figure 11. Orders displayed on this page are sorted chronologically based on their placement date, facilitating a systematic approach for sellers to fulfill orders in the sequence they were received. Upon successful delivery to the buyer, sellers possess the capability to update the order status to "Delivered." This intuitive system empowers sellers to efficiently manage their order fulfillment process, ensuring timely updates and clear communication throughout the transaction lifecycle.

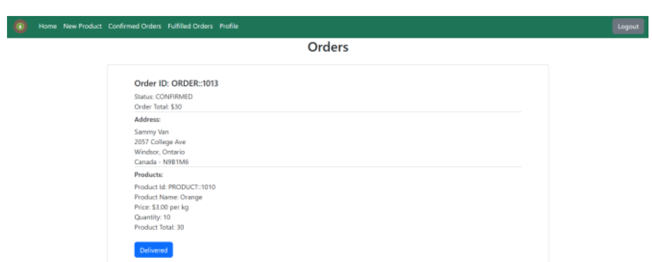

Figure 11. Confirmed orders of the seller

The "Fulfilled Orders" tab serves as a comprehensive record of all orders successfully completed by the seller, providing them with convenient access to their order fulfillment history as shown in Figure 12. This feature enables sellers to efficiently track and review past transactions, fostering improved organization and accountability within their business operations.

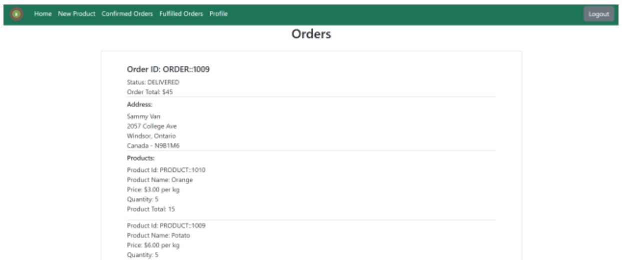

Figure 12. Fulfilled Orders page of the seller.

Sellers have the capability to seamlessly integrate new products into their inventory by completing essential details such as product name, description, quantity, and per unit price as shown in Figure 13. Additionally, they can enrich product listings by uploading corresponding images, which are securely stored within Google Cloud Platform (GCP) buckets for efficient asset management. This streamlined process empowers sellers to present their products with compelling visuals while ensuring all pertinent information is readily available to potential buyers, thereby enhancing the overall shopping experience and promoting sales growth.

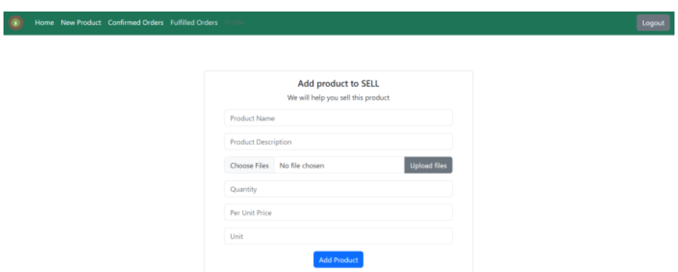

Figure 13. Add the products page of the seller.

The seller home page serves as a centralized hub for managing inventory, displaying a comprehensive list of all products available for sale as shown in Figure 14. From this interface, sellers can effortlessly update product details and adjust available quantities as needed, ensuring accurate and up-to-date listings.

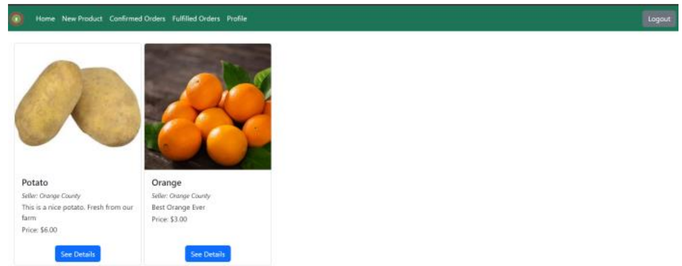

Figure 14. Seller home page

Figure 15 shows the model for training a deep learning model, particularly for Image classification based on ripeness. The dataset comprises 643 training images and 161 validation images, with each image containing multiple tomatoes. To effectively label each image, a majority-based approach is employed, where the class assigned to an image is determined by the majority class of tomatoes within it. For instance, if 9 out of 10 tomatoes in an image are ripe, the image is labeled as "ripened." This method ensures that each image is accurately categorized based on the dominant class of tomatoes present, enhancing the quality of the dataset for training and evaluation purposes in tomato classification tasks. After loading the image, it defines functions to read labels from text files associated with each image and preprocess the images to resize them to a target size, pad them if necessary, and normalize their pixel values. The images and labels are loaded and organized into arrays for both the training and validation sets.

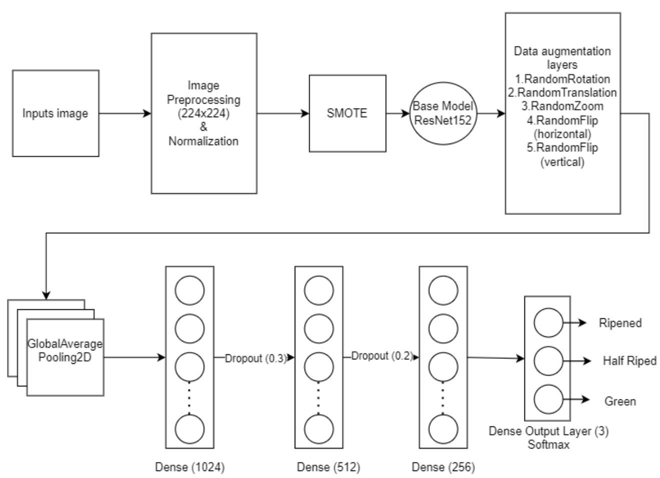

Figure 15. Model for image classification.

After preparing the dataset, the code applies the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance in the training data. SMOTE is applied to training images reshaped into a format compatible with the SMOTE algorithm. The resampled images and labels are then reshaped back to their original dimensions. Subsequently, a pre-trained ResNet152 model is loaded, and additional data augmentation layers, including random rotation, translation, zoom, and flipping, are added on top of the base model. The model is then compiled and trained over 100 epochs, using the resampled training data, with class weights computed to handle class imbalance. Finally, the model is evaluated on the validation dataset, and the validation loss and accuracy are printed and outlined in Figure 16. This code serves as a comprehensive pipeline for preparing a dataset, addressing class imbalance, and training a deep learning model for object classification.

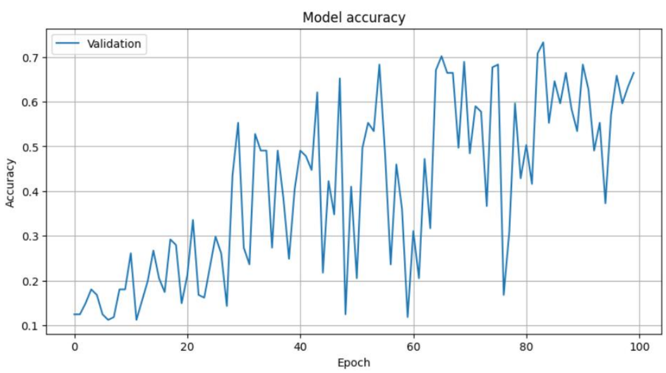

Figure 16. Validation accuracy for each epoch

We can observe that the model accuracy is gradually increasing over the duration of epochs in figure 16.

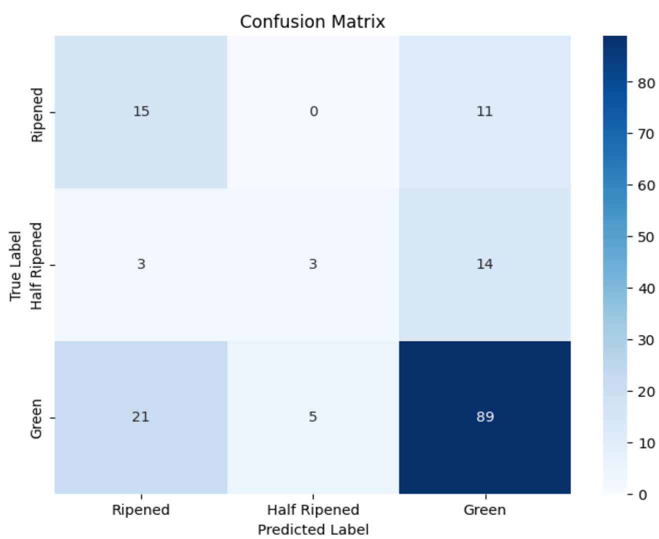

Figure 17. Confusion matrix for the trained model

After rigorous evaluation, the model utilizing ResNet152 transfer learning achieved a commendable validation accuracy of 66.4%, surpassing the performance of the ResNet50 model, which attained an accuracy of 54.6%. Consequently, our decision to adopt ResNet152 as the preferred model is founded on its superior accuracy and effectiveness in capturing intricate features within the dataset which is shown in figure 17.

We have developed a model capable of both classification and price segmentation specifically tailored for tomatoes, with plans to extend this system to accommodate other products in the future. When a seller uploads an image of a tomato, our system intelligently determines its ripeness using a pre-trained model based on transfer learning with ResNet architecture. Subsequently, if the tomato is classified as unripe, semi-ripe, or ripe, our system dynamically recommends an appropriate price based on the quality of the product. This is shown in figure 18 For instance, unripe tomatoes may be suggested at a reduced price due to their lower market value, while semi-ripe tomatoes, known for their extended shelf life, may warrant a higher price point. Conversely, ripe tomatoes, with a shorter lifespan, may prompt a different pricing recommendation. This functionality is facilitated through a Flask API, seamlessly integrating image classification with real-time price segmentation to optimize sales strategies and enhance overall efficiency for sellers.

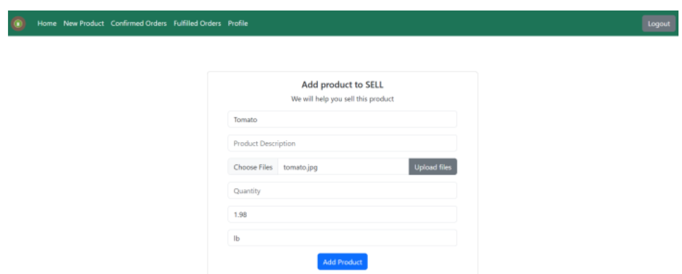

Figure 18. Price recommendation based on price segmentation.

6. LIMITATIONS OR CHALLENGES

Implementing password less authentication offers improved user experience and security benefits yet poses significant challenges. Ensuring the robustness of token generation and validation is paramount, as any vulnerabilities could compromise the entire authentication system. Managing and securing JWT tokens, typically stored in cookies, adds complexity, as mishandling or improper storage may expose sensitive user data. Moreover, reliance on email for token dispatch introduces further complexity and potential authentication delays, requiring a robust email delivery system and accounting for user email accessibility issues. Addressing the constraints imposed by a limited dataset size proved to be a formidable challenge, given its tendency to constrain the model's capacity for robust generalization. The pursuit of optimal performance through meticulous fine-tuning necessitated an exhaustive journey marked by iterative experimentation and refinement of hyperparameters. While our efforts culminated in achieving an acceptable level of accuracy, it remains evident that there exists untapped potential for further enhancement in the model's performance.

7.FUTURE WORK

The future trajectory of "Produce Sync" involves a multifaceted approach to enhance its capabilities and impact in agricultural settings. Additionally, efforts will focus on expanding the product range to encompass a wider variety of crops, integrating IoT devices for real-time monitoring, ensuring scalability and global adaptability, conducting user acceptance studies, assessing sustainability impacts. These endeavors collectively aim to revolutionize agricultural practices, promote sustainability, and optimize the agricultural supply chain for the benefit of farmers, consumers, and the environment.

8. CONCLUSION

In summary, the "Produce Sync" initiative has demonstrated an example of the impact that technology can have on the agricultural sector. This project addresses the significant problem of food waste by integrating computer vision, adaptive pricing, and the capability to accurately determine the location of produce. Not only does it enhance the efficiency of the food supply chain from farms to consumers, but it also presents an attractive strategy for implementing this on a global scale. This effort suggests an advancement in enhancing the sustainability of agriculture and ensuring our ability to satisfy the global food demands without causing harm to our environment.

The research conducted with "Produce Sync" demonstrates that the effective utilization of technology can address significant challenges in the cultivation and distribution of food. This project presents significant concepts that can be utilized and expanded upon by other professionals in the agricultural domain, with the objective of enhancing efficiency, reducing waste, and promoting equity in food management practices.

As we anticipate the future, the accomplishments of "Produce Sync" should motivate us to continue seeking new and superior methods of utilizing technology in agriculture. There exists a lot of room for further enhancements in the realm of food production and delivery, with "Produce Sync" serving as just an initial step towards this objective. Through continuous innovation and collaborative efforts, we can create a future in which global access to fresh, sustainable food is achieved, and our agricultural methods actively contribute to the well-being of the planet by reducing food waste.

10.REFERENCES

[1] H. Onyeaka, P. Tamasiga, U. M. Nwauzoma, T. Miri, U. C. Juliet, O. Nwaiwu, and A. A. Akinsemolu, "Using Artificial Intelligence to Tackle Food Waste and Enhance the Circular Economy: Maximising Resource Efficiency and Minimising Environmental Impact: A Review," Sustainability, vol. 15, no. 13, p. 10482, 2023. [Online]. Available: https://doi.org/10.3390/su151310482

[2] M.P. Arakeri, L. Lakshmanab, "Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture Industry," in Procedia Computer Science, vol. 79, pp. 426-433, 2016.

[3] H. Charles J. Godfray et al., Food Security: The Challenge of Feeding 9 billion People.Science 327,812-818, 2010, DOI:10.1126/science.1185383

[4] Vijay Kakani, Van Huan Nguyen, Basivi Praveen Kumar, Hakil Kim, Visweswara Rao Pasupuleti, A critical review on computer vision and artificial intelligence in food industry, Journal of Agriculture and Food Research, Volume 2, 2020, 100033, ISSN 2666-1543, https://doi.org/10.1016/j.jafr.2020.100033.

[5] LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015). https://doi.org/10.1038/nature14539

[6] A. Payasi and G.G Sanwal, Journal of Food Biochemistry, 34:679-710, 2010, https://doi.org/10.1111/j.1745-4514.2009.00307.x

[7] Borawar, L., Kaur, R. (2023). ResNet: Solving Vanishing Gradient in Deep Networks. In: Mahapatra, R.P., Peddoju, S.K., Roy, S., Parwekar, P. (eds) Proceedings of International Conference on Recent Trends in Computing. Lecture Notes in Networks and Systems, vol six hundred. Springer, Singapore. https://doi.org/10.1007/978-981-19-8825-7_21

[8] Agarwal, S., Rajan, K.S. Performance analysis of MongoDB versus PostGIS/PostGreSQL databases for line intersection and point containment spatial queries. Spat. Inf. Res. 24, 671–677, 2016, https://doi.org/10.1007/s41324-016-0059-1

[9] Marc Ruef, "WebAuthn - The Future of Web Authentication," April 2018.

[10] J. Gramila, "Building a WebAuthn Application with Java," Okta Developer, 26-Apr-2022.