International Journal of Engineering and Techniques - Volume 12 Issue 2, March-April - 2026

AI-Powered Food Label Analyzer: Enhancing Consumer Awareness Using OCR and Machine learning

K. Jagan Mohan Reddy
Department of Artificial Intelligence and Data Science,
Nalla Malla Reddy Engineering College
Hyderabad, Telangana, India
JaganMohanReddy1703@gmail.com

S. Ram Chandra Reddy
Department of Artificial Intelligence and Machine Learning,
Nalla Malla Reddy Engineering College
Divya Nagar, Hyderabad
rcreddy79@gmail.com

Ravi Bukya
Department of Artificial Intelligence and Machine Learning
Nalla Malla Reddy Engineering College
Divya Nagar, Hyderabad
drravibukya@gmail.com

Bikki Bharath
Department of Artificial Intelligence and Data Science,
Nalla Malla Reddy Engineering College
Hyderabad, Telangana, India
bikkibharathgoud23@gmail.com

S.G. Sriniketh
Department of Artificial Intelligence and Data Science
Nalla Malla Reddy Engineering College
Hyderabad, Telangana, India
srinikethsannidhiraj@gmail.com

Pambi Rajesh
Department of Artificial Intelligence and Data Science
Nalla Malla Reddy Engineering College
Hyderabad, Telangana, India
rajeshpambi18@gmail.com

***Abstract*— Although to have a healthy lifestyle, it is crucial to understand what is written on food product labels, a majority of consumers find it hard to decipher nutritional data and ingredient lists because of the complicated terms used and their ignorance. In this paper, the author will introduce an AI-based food label analyser which combines both the Optical Character Recognition (OCR) and barcodes scanning methods to deliver a practical and convenient solution to the analysis of packaged food products. The system proposed allows users to scan a product barcode or take an image of a food label. OCR is utilized to capture unstructured textual data like ingredients, nutritional values and additives to product packaging and barcode scanning is used to quickly access structured product information as it is stored in readily accessible databases. The information extracted are then processed and analysed with a machine learning-based classification technique to assess the overall nutritional quality of the product.**

This system is aimed at the detection of the most crucial health indicators, such as the sugar level, the fat content, the concentration of sodium and the presence of artificial additives or preservatives. Using this analysis, the application assigns food items to various levels of health and offers valuable insights to help users make healthy food choices. It is created with Flutter to be deployed on all platforms and uses Google ML Kit to detect barcodes and OCR in real-time. Experimental analysis shows that the system is very accurate in extracting text as well as reliable in barcode recognition to different conditions. OCR and barcode scanning together is more flexible, data is more reliable, and is more usable in real-life situations. On the whole, the suggested system will help spread health awareness by making the interpretation of food labels easier and making it possible to analyse nutrition in real-time and intelligently, using an accessible mobile platform.

INTRODUCTION

The consumption of packaged and processed food products has increased tremendously over the last few years because of changing lifestyles and urbanization. Although food packaging gives the necessary information like ingredients, nutritional values and expiry dates, most consumers cannot effectively interpret this information. Complex terminologies, small fonts, and nutritional ignorance can easily result in poor informed choices of foods, which are susceptible to various health complications including obesity, diabetes, and cardiovascular diseases.

As the development of artificial intelligence and mobile technologies is rapidly evolving, it is possible to come up with intelligent systems capable of helping users to learn more about food products with minimal effort. Specifically, machine learning and computer vision have demonstrated a lot of potential in text recognition and analysis of text in images. The conversion of printed text to machine-readable form is provided by the Optical Character Recognition (OCR), and the rapid access to the structured product information stored in databases is made possible through barcode scanning. The paper introduces an AI powered food label scanner that combines OCR and barcodes scanning to offer a complete food product analysis solution. The system can enable a user to take a picture of a food label or scan a barcode with a mobile device. OCR module takes out unstructured textual data like ingredients and nutritional values right off product packaging, and the barcode module finds product-specific information effectively.

The data obtained is then further processed and analysed with the help of machine learning to assess the nutritional value of the product. The sugar content, fat, sodium level, and the use of artificial additives are also important parameters that are used to categorize food items as healthy, moderate and unhealthy. The classification aids users in making wise choices regarding diet and is also useful in creating awareness regarding food habits. It is created with Flutter, which allows cross-platform functionality, and Google ML Kit, which provides real-time OCR and barcode scanning. The system will be user friendly, efficient and able to give real time results. The proposed solution will improve the use of different technologies in one platform, increase the accuracy, and solve the shortcomings of the current food label analysis systems. In general, the purpose of this work is to fill the gap between the complicated nutritional information and the understanding of the user, to offer a smart, easy-to-use, and practical tool to be used on a daily basis.

RELATED WORK

The issue of food labelling awareness and nutritional analysis has been extensively discussed over the last few years because of more health-conscious food consumption and awareness of food safety. Conventional methods were based on the time-

consuming manual examination of food labels and fixed nutritional databases that are subject to human mistakes and errors. Initial studies were done on the application of barcode scanning as a major mode of food identification. The systems based on barcodes enable users to access product data like nutritional value and ingredients by scanning the Universal Product Code (UPC) with a smartphone camera. They have found extensive use in diet-tracking applications, offering fast and accurate access to product structured data. Their performance, however, is limited to the availability and correctness of external databases that do not necessarily have up-to-date or location-specific products. In order to solve the shortcomings of the barcode-only systems, scientists have considered the application of Optical Character Recognition (OCR) as a method of retrieving textual information on the surface of food packaging. OCR systems allow identifying ingredients, allergens, and nutritional values of the labels automatically. As an example, recent research has suggested OCR-based allergen detection systems, which identify harmful ingredients and offer alternative food options to users with dietary limitations. Moreover, AI-based OCR technologies have been used in food traceability and supply chain management, enhancing transparency and minimizing human error when verifying labels.

Irrespective of these developments, OCR-based solutions in practice have various challenges. The packaging of food may have complicated layouts, font differences, rounded objects and lack of light, and these factors can have considerable impact on the accuracy of text recognition. Research assessing the performance of OCR in food labelling indicates that multilingual text, dense formatting, and image noise are among the challenges that complicate accurate extraction of the text in food labels. These drawbacks suggest that strong preprocessing and hybrid strategies should be applied to enhance the performance of the systems.

Recent trends were geared towards incorporating machine learning and artificial intelligence methods with the OCR and barcode systems. Extracted nutritional data is analysed using machine learning models and classify food products according to health parameters sugar, fat, and sodium content. Recommendation engines are also used in some systems to propose healthier options, or customized diets. Moreover, large-scale food composition databases have been built to facilitate this kind of application through providing formalized nutritional details of a broad assortment of products. Even though a lot of progress has been achieved, most of the existing systems concentrate on either text extraction by using OCR or product identification by barcode on its own. Only a handful of solutions provide a comprehensive system that allows data mining and data acquisition to coexist as one system with user-friendly mobile apps.

We overcome these limitations in this work, where we propose an AI-based food label analyser, combining both OCR and barcode scanning and machine learning-based analysis into a single system. This integrated solution enhances accuracy, usability, and offers a more holistic solution to real time food label interpretation and nutritional evaluation.

METHODOLOGY

Dataset sources

The system uses various data sources such as:

- Food nutrition databases that are publicly accessible like packaged food repositories with a record of nutritional values (calories, sugar, fat, protein, etc.).
- The data on products gathered with the help of barcode-based data sets, including the name of the item, brand name, and list of ingredients.
- Live information scraped off food labels with the OCR so that the system can operate even with products that are not available in an online database.
- Knowledge bases of ingredients so as to recognize harmful ingredients like artificial preservatives, colorants, and additives.

These data sets are useful in classifying and analysing food items in respect to their nutritional value.

Data Preprocessing

The datasets were cleaned up before training the model to improve data quality and consistency. The following steps were applied:

- De-noising of OCR-extracted text, such as special symbols and formatting mistakes.
- Text to standard form (e.g., text to lowercase, text to unit normal (e.g. grams, milligrams)).
- Listing of ingredients that are tokenized to find single ingredient to analyse.
- Dealing with ambiguous expression of nutritional values (e.g. sugar:10g vs. 10 grams sugar).
- OCR error detection and correction by use of rule based or dictionary-based methods.

These preprocessing steps are necessary to improve the model performance and reduce noise in the data.

OCR and Barcode Detection

The system uses Google ML Kit for**:**

- Textual information on food labels is extracted and converted into textual data through the application of Optical Character Recognition (OCR) on the image of food labels obtained with the help of a mobile camera.
- The OCR module is recognized with the help of Google ML Kit to detect and recognize texts in real time.
- It pulls out the important information like the ingredients, nutritional facts and the descriptions of the products.
- The preprocessing of the images with noise reduction, image resizing and text alignment is used in order to enhance OCR accuracy.
- The system also manages font differences in font styles, font sizes, font orientations, and light events that occur in the real-life packaging.
- Product identification is done by scanning barcodes to identify the products fast and accurately.

MACHINE LEARNING MODEL

The analysis and classification of food products according to their nutritional content are carried out with the help of a supervised machine learning approach. The features used as input in the model include sugar level, fat content, sodium level and the presence of additives. The extracted features on OCR and barcode modules are processed into structured numerical format that is used to process the models.

The methods of feature selection are used to determine the most pertinent nutritional parameters that may affect the food quality. The data is separated into training and testing data to analyse the model performance. Data normalization

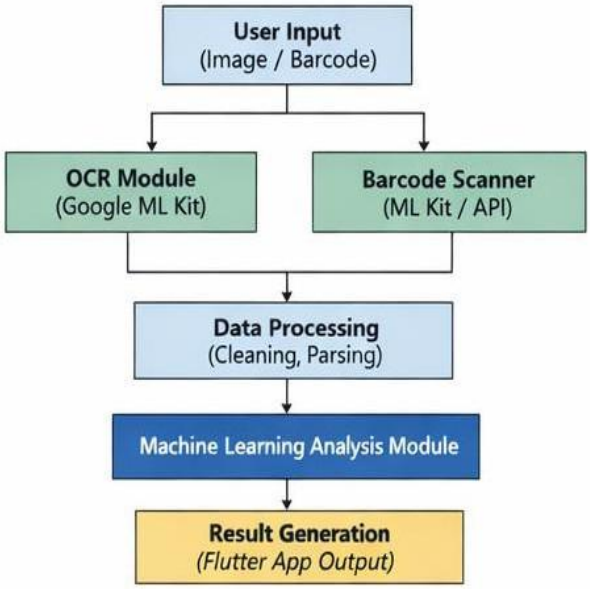
algorithms are used to make the input features uniformly scale. Food items are categorized using a classification algorithm (e.g., Decision Tree / Random Forest / Logistic Regression).

SYSTEM ARCHITECTURE

The proposed AI-powered food label analyser has its system architecture implemented as a modular and scalable architecture that combines image processing, data extraction, and intelligent analysis into a single pipeline. The architecture comprises various interrelated parts that collaborate to give real-time food label reading and nutrition evaluation.

- Input Layer-This is initiated at the input layer whereby the user enters the data by scanning a product barcode with a mobile device or by taking an image of food label. This two-input system offers flexibility and applicability to various real-life situations.
- Data Acquisition Layer-The input captured is then sent to the data acquisition layer comprising of the OCR module and barcode scanning module. The OCR module detects text-based information like ingredients and nutritional values on the label pictures and the barcode module reads the structured information on products based on the available databases. These modules are autonomous but combined to offer complementary services.
- Processing Layer-After gathering data, the information obtained is transmitted to the data processing layer where preprocessing methods are used. This layer carries out activities like removal of noise, normalization of texts, unit conversion, and arrangement of extracted data into significant fields. This is an essential step of enhancing the quality and consistency of the input data prior to analysis.
- Machine Learning Layer-The processed data will then be sent to the machine learning layer, which will analyse the nutritional content of the product. Some of the critical health indicators that are assessed by this state are sugar, fat, sodium, and additives, and the food item is categorized into preset groups such as healthy, moderate, or unhealthy. Machine learning can be used to make intelligent decisions and improve the overall system functionality.
- Application Layer-Lastly, the application layer delivers the results, and it is provided in form of a user-friendly mobile interface. The system shows the results of the analysed data, nutritional knowledge, and health suggestions in a clear and simple manner.

Overall, the architecture ensures efficient data flow, modular design, and real-time performance. The combination of OCR, barcode scanning, and machine learning in one platform enhances accuracy, flexibility, and usability, and makes the system an efficient approach to implementation.



TECHNOLOGIES USED

The proposed system exploits both the use of modern technologies to bring about effectiveness in terms of data extraction, processing and analysis. The selection of these technologies is well-considered to make sure they work in real-time and are accurate and cross-platform.

Flutter is used for developing the mobile application interface. It gives it an effective and flexible structure of developing cross-platform applications to allow the system to run smoothly on both Android and iOS platforms. Flutter is also fast in creating UI and delivering a seamless user experience.

Google ML Kit is also used to implement OCR and barcode scanning capabilities. It has on-device machine learning functionality, enabling real-time text recognition and barcode detection without needing to be connected to the internet at all times. This enhances the performance **of** the system in terms of speed, reliability and privacy.

The nutritional data and food product classification are analysed using Machine Learning Algorithms. Models like Decision Trees, Random Forest, or Logistic Regression which are lightweight models are used to make sure that they can be used efficiently on mobile devices and still achieve acceptable accuracy. These models allow one to automatically assess the quality of food in terms of features extracted.

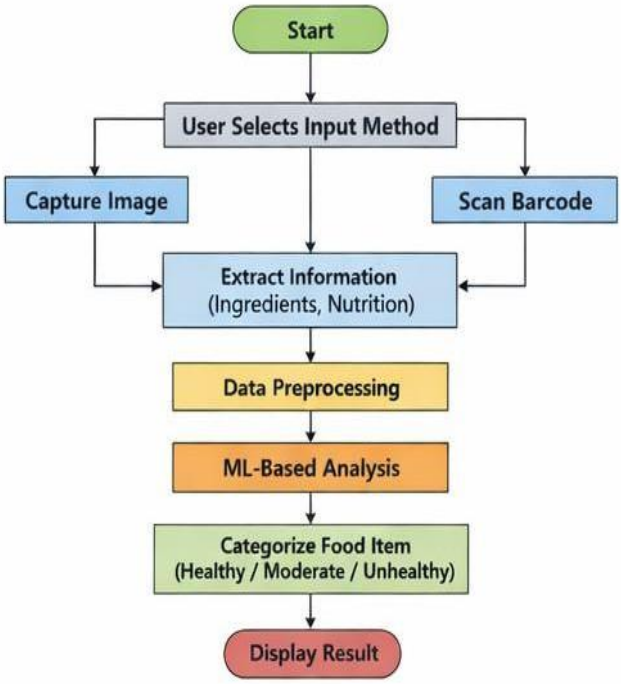
Before OCR Processing, Image Processing Techniques are used to improve the quality of images taken. Photoshop tricks like resizing, reduction of noise, and adjustment of contrast are useful in enhancing the accuracy of text recognition, particularly in different lighting situations.

Information about the products can be obtained through Database and API Integration to find information on products by entering barcodes. These third-party data sources will present well-organized nutrition information, enhancing the accuracy and thoroughness of the analysis.

Programming Languages and Tools Programming Languages and Tools include Dart (to develop a Flutter application) and the libraries that support the application logic and connect various modules.

In general, a strong, efficient, and scalable food label analyser can be developed due to the combination of these technologies. The combination of mobile development frameworks, machine learning, and computer vision algorithms guarantees

that the system achieves the correct results at real-time performance and user comfort.

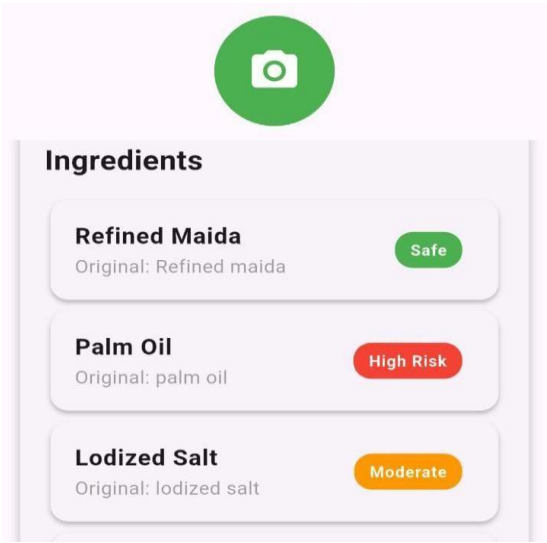


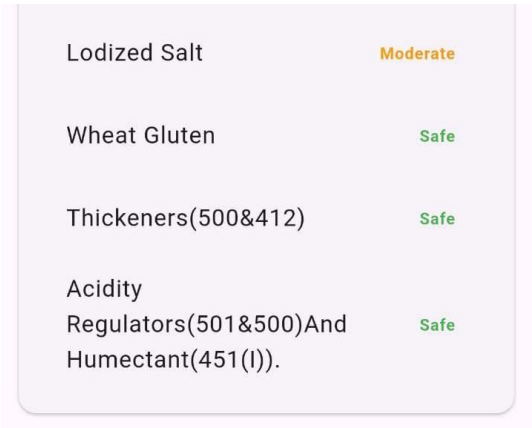
RESULTS AND DISCUSSION

The AI-powered food label analyser was tested according to its text extractions, barcode identification, and nutritional analysis. The system was experimented with on a range of packaged food products in various conditions such as lighting, image quality and label design. There was good consistency in the OCR module in extracting textual data like ingredients and nutritional value. It had a high precision with distinctly printed labels, and smaller errors were noticed when the quality of the image was low, the layout was complicated, or the surface of the packaging was curved. The combination of preprocessing methods contributed to enhancing the text recognition performance to a considerable degree.

Barcode scanning module offered quick and precise product identification. It was able to retrieve structured information about the product with a low processing time when the barcode was visible. Barcode scanning was more effective and reliable than OCR and its performance was determined by the presence of product data in the database. The machine learning model successfully processed the obtained nutritional data and put food items into healthy, moderate, and unhealthy categories. The outcomes of the classification were in line with the general nutritional guidelines especially in determining products rich in sugar, fat and sodium. The application of important nutritional parameters helped to obtain significant and interpretable results and barcode scanning were used together which enhanced the overall system performance. Although barcode scanning was quick in accessing structured data, OCR ensured that the system was not rendered useless in the event that barcode data was not available. The hybrid solution improved the reliability and usability of the systems in the real world.

In general, the system was found to be efficient, respondent in real-time and applicable in practice. Nevertheless, issues like poor OCR in low condition and reliance on external databases to access barcode information are areas to improve on.





CONCLUSION AND FUTURE SCOPE

This paper introduced an AI-driven food label analyser that combines OCR, barcode scanning, and machine learning methods to make it easier to understand information about food products. With a mobile app, the system lets users quickly get nutritional information from packaged foods and analyse it.

Combining OCR and barcode scanning lets you quickly get structured product information from food labels and unstructured textual data from food labels. The machine learning model makes the system even better by looking at the nutritional value of food and sorting it into groups based on health factors.

The tests show that the proposed system can give users accurate and useful information that helps them make smart choices about what to eat. The app is easy to use, works quickly, and can be used in real time, making it a good choice for promoting health awareness.

There are many ways to make the proposed system even better so that it works better and can be used more widely. One possible improvement is to use more advanced deep learning models to make OCR more accurate, especially for images that

are hard to read or of low quality. You can also connect the system to big cloud-based food databases to get more complete and up-to-date information about products. Adding personalised diet suggestions based on a user's health profile can also make the app much more useful.

Adding support for more than one language can make the system easier for more people to use. Real-time voice assistance and wearable device integration can also make it easier for users to interact with the system. These changes will make the system smarter, more flexible, and more focused on the needs of users

ACKNOWLEDGMENT

This project titled "AI-Powered Food Label Analayzer" was carried out with guidance and support from several people.

We would like to thank Dr. Rama Chandra Reddy for guiding us during the project work and for clarifying concepts related to artificial intelligence and its use in healthcare applications. His suggestions helped us complete the project successfully.

We also thank the Department of Computer Science and Engineering for providing the required facilities to work on this project.

We are thankful to our friends and family members for their support during the project period.

REFERENCES

1. Parkavi,"Android App for Food Allergen Detection,"IEEE,2023.
2. K.Azzimani,"AI-Based Personalized Nutrition,"Information Fusion,vol.78,pp.1-12,2022.
3. S.-C.Huang and W.-C.Chiang,"An Image-Based AI Nutrition Analysis Platform for Tray Foods,"2023.
4. D.Khairani,"Hala Ingredient Detection Via OCR and Named Entity Recognition,"IEEE,2022.
5. S.P.Siddique Ibrahim,"Nutritional Claim verification Using Ai,"2024.
6. S.Xiong and G.Bai,"Food Safety Risk Detection Using Artificial Intelligence,"IEEE,2023.
7. P.Kwon,"Smart Nutrition Application for Children,"IEEE,2023.
8. H.A.AI-Manwari,"Smart Food Analyzer System,"2023.
9. D.Ilango,"Toxin Detection in Food Using AI,"IEEE,2024.
10. S.Kayalvizhi,"Detection of Unsafe Ingredients in Packaged Food,"IEEE,2023.
11. N.Prabha,"Nutriknow:AI-Based Ingredient Analyzer,"IJISRT,2025
12. S.Lodha,"Nutriscan:AI-Based Ingredient Detection and Evaluation,"IJERT,2025.
13. Y.Shah,"Nutrition Table and Ingredient Recognition Using Ai,"IJRASET,2025.
14. P,Ma et al.,"UMDFood:Vision-Language Models for Food Composition Compilation,"arXiv preprint arXiv:2306.01747,2023.
15. Bukya, Ravi, G. Madhu Mohan, and M. Kumar Swamy. "Artificial Intelligence role in optimizing electric vehicle charging patterns reduce costs and improve overall efficiency: A review." Journal of Engineering, Management and Information Technology 2.3 (2024): 129-138.
16. R.L.Bailey,"ArtificialinNutritionResearch,"PNAS Nexus,2024