iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** XI **Month of publication:** November 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.56852

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue XI Nov 2023- Available at www.ijraset.com*

# Delving Deep into NutriScan: Automated Nutrition Table Extraction and Ingredient Recognition

Yaksh Shah[1], Nandan Jariwala[2], Bhakti Kachhia[3], Prachi Shah[4]
*Birla Vishvakarma Mahavidyalaya, Mota Bazaar, Anand-388120, India*

***Abstract:*** ***Our paper introduces an innovative solution for automating the detection of nutrition tables and recognition of ingredients on packaged food products, addressing the growing demand for tools that facilitate the extraction of essential dietary information from food packaging. The increasing interest in healthy eating and dietary awareness underscores the need for precise and efficient tools to assist consumers and health-conscious individuals in making well-informed food choices. Harnessing state-of-the-art technology, we enhance the accuracy of nutrition table and ingredient detection through fine-tuning the EfficientDet model. We further extract textual content from these tables using PaddleOCR a sophisticated optical character recognition tool. Additionally, we employ regular expressions to meticulously capture crucial nutritional details, including salt or sodium content, carbohydrates, total fat, saturated fat, trans fat, protein values, additives, and allergens found in the product. To empower users in their dietary decision-making, we introduce a dynamic user preference system. Users can tailor their dietary requirements, specifying allergies and preferences for sugar, saturated fat, total fat, or salt content. Our system then evaluates food products and offers personalized recommendations including the nutri-score, indicating their alignment with the user's health goals. Extensive experimentation and evaluation confirm the high accuracy and efficiency of our method in both nutrition table and ingredient recognition. These paper underscores the transformative potential of our approach, which can revolutionize the accessibility of dietary information, encourage healthier eating habits, and promote dietary transparency.***

***Keywords: Nutrition and Ingredient Detection, Optical Character Recognition, user-preferences, nutri-score, dietary decision-making***

## I. INTRODUCTION

In today's health-conscious world, where the quest for nutritious and balanced diets is on the rise, access to accurate dietary information has become paramount. Consumers and individuals striving for healthier eating habits often turn to the nutritional content provided on food packaging as a primary source of guidance. However, the process of extracting pertinent dietary details from food labels and packaging can be time-consuming and error-prone when performed manually. This challenge has spurred the need for innovative solutions that can automate the identification of nutrition tables and recognition of ingredients, facilitating more informed food choices and enhancing dietary transparency. Our paper introduces a pioneering approach to address this challenge. We present an automated system that leverages cutting-edge technology to enhance the precision of nutrition table and ingredient detection from packaged food products. The increasing interest in dietary awareness and the pursuit of healthier lifestyles underscores the significance of this endeavor. In this digital age, we capitalize on the capabilities of advanced machine learning techniques and optical character recognition to revolutionize the way individuals access dietary information. By automating the extraction of crucial dietary details from food packaging, we offer consumers and health-conscious individuals a convenient and efficient means to make informed dietary choices. Our solution not only streamlines the process but also empowers users by allowing them to customize their dietary preferences, tailoring recommendations to align with their health objectives. This paper presents a comprehensive overview of our approach, detailing the methodology and technology behind our automated system. We also provide evidence of the system's accuracy and efficiency through extensive experimentation and evaluation, highlighting its potential to usher in a new era of dietary transparency and healthy eating habits. Through this work, we aim to contribute to the advancement of technology in the field of nutrition and empower individuals to make informed, health-conscious choices in their daily food consumption

## II. LITERATURE SURVEY

In recent years, the design of efficient neural network architectures for computer vision tasks, such as object detection, has gained paramount importance. A pivotal contribution in this domain is the paper EfficientDet by Mingxing Tan et al.,[1], which systematically investigates neural network architecture choices to enhance efficiency in object detection. The key optimizations proposed in this work include, Weighted Bi-directional Feature Pyramid Network (BiFPN) this novel feature pyramid network (FPN)

©IJRASET: All Rights are Reserved 

[2] design facilitates rapid multiscale feature fusion, simplifying the process and improving efficiency in object detection. Compound Scaling Method this method uniformly scales the resolution, depth, and width for all components of the network, including the backbone, feature network, and box/class prediction networks, simultaneously.

The culmination of these optimizations, coupled with improved backbones, results in a family of object detectors known as EfficientDet. These detectors consistently outperform prior approaches in terms of efficiency across a broad range of resource constraints. Remarkably, with single model and a single scale, EfficientDet-D7 attains a state-of-the-art 55.1 AP on the COCO test-dev dataset, all while using only 77 million parameters and 410 billion FLOPs. This represents a significant advancement as it is 4 to 9 times smaller and consumes 13 to 42 times fewer FLOPs compared to previous detectors [1-3].

For OCR purpose we use PP-OCR by Yuning Du et al.,[4] which is an optical character recognition (OCR) system that is designed to be both accurate and efficient. It is based on a simple segmentation network that integrates feature extraction and sequence modeling. PP-OCR uses Differentiable Binarization (DB) as its text detector, which is a lightweight and effective approach. PP-OCR has been shown to achieve state-of-the-art results on a variety of benchmark datasets, including ICDAR2015 and COCO-Text. It has also been successfully deployed in number of real-world applications, such as mobile OCR apps and document scanners [4].

The FSA Nutri-Score is a nutrition label that ranks the nutritional quality of foods and beverages from A (most healthy) to E (least healthy). It was developed by the UK Food Standards Agency (FSA) and is based on a nutrient profiling system that takes into account the levels of energy, saturated fat, sugar, salt, fiber, protein, fruits, vegetables, legumes, nuts, and rapeseed. [5]

## III. METHODOLOGY

### *A. Data Gathering and Augmentation*

In our quest to assemble a robust and diverse dataset for our research, we embarked on a photographic journey through various retail supermarkets. With our cameras, we captured many different types of food packets. With great care and precision, we meticulously annotated these images, demarcating the essence of the food packaging using bounding boxes. Through data augmentation, we infused our dataset with a burst of creativity and diversity. We blurred bounding boxes, allowing our model to engage with the intricacies of real-world scenarios. The colors of our images were enriched through saturation techniques. We added twists and turns by skillfully rotating our images, ensuring our dataset covered a broad spectrum of visual perspectives. After all the steps we had a dataset of almost 1500 images to be utilized for training purpose.

In addition to formation of our image dataset, we recognized the pivotal importance of providing users with valuable insights into the ingredients of the packaged food items they capture. To this end, we turned to the treasure trove of knowledge that is Wikipedia [6]. Drawing from the wealth of information available on Wikipedia, we gathered data about food additives commonly found in the ingredients of various packaged food products. When users interact with our system and capture a food item, this data repository empowers us to offer them a comprehensive overview of the additives present in the ingredients, enhancing their understanding of the products they encounter. These unique integration of data sources not only enriches the user experience but also showcases our commitment to providing informative and useful insights to our audience.

### *B. Model Architecture*

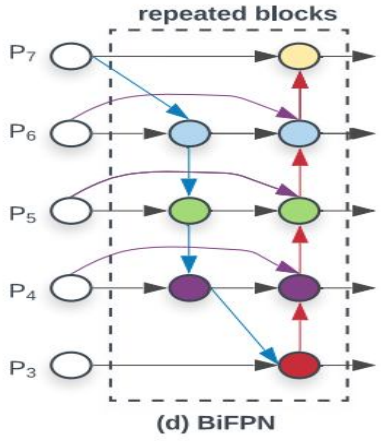



Figure.1 Bidirectional Feature Pyramid Network

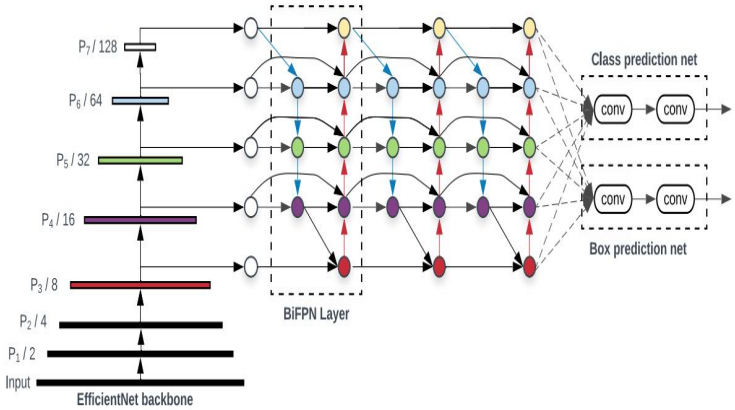



Figure.2 EfficientDet Architecture

EfficientDet [1] is a family of object detection models that are designed to be both accurate and efficient. The models are based on the EfficientNet [2] backbone, which is a family of convolutional neural networks that are designed to achieve state-of-the-art accuracy with minimal computational resources. In addition to the EfficientNet backbone, It also uses several other optimization techniques to improve efficiency, including BiFPN (Bi-directional Feature Pyramid Network) which is a feature fusion network that allows for more efficient and effective multiscale feature fusion and compound scaling method that uniformly scales the resolution, depth, and width of all backbones, feature networks, and box/class prediction networks at the same time. This allows the models to be scaled up or down without sacrificing accuracy or efficiency.

EfficientDet models consist of three main components:

1) *Backbone*: The backbone is responsible for extracting features from the input image. EfficientDet uses the EfficientNet backbone, which is a family of convolutional neural networks that are designed to achieve state-of-the-art accuracy with minimal computational resources.
2) *Feature network*: The feature network is responsible for fusing features from different levels of the backbone to produce a more informative feature representation. EfficientDet uses the BiFPN (Bi-directional Feature Pyramid Network) feature fusion network, which allows for more efficient and effective multiscale feature fusion.
3) *Box/class prediction network*: The box/class prediction network is responsible for predicting the bounding boxes and class labels of the objects in the image. EfficientDet uses a simple head network that consists of a few convolutional layers and a fully connected layer.

### C. Model Training

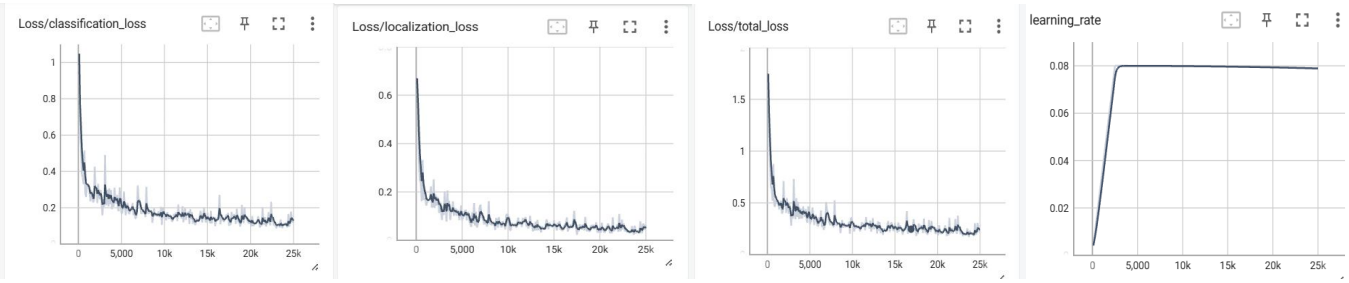
In the pursuit of achieving optimal performance and accuracy for our specific application, we embarked on an extensive model training journey. Our choice of architecture, the EfficientDet model, has already been introduced for its remarkable efficiency and effectiveness. However, to adapt it for recognizing Nutrition table and ingredients, we fine-tuned this model on our meticulously curated dataset. The fine-tuning process, a pivotal phase of our model development, consisted of training the EfficientDet model for an extensive 25,000 steps. To facilitate this, we used TensorFlow Object Detection (TFOD) API, a powerful tool for custom object detection tasks.

### D. Model Testing




This is the result we get after testing the model on the live images captured by the camera of our phone.

### E. Model Evaluation




These are the final evaluation graphs based on the COCO Metrics that are generated using tensor-board.

## IV. APPLICATION ARCHITECTURE

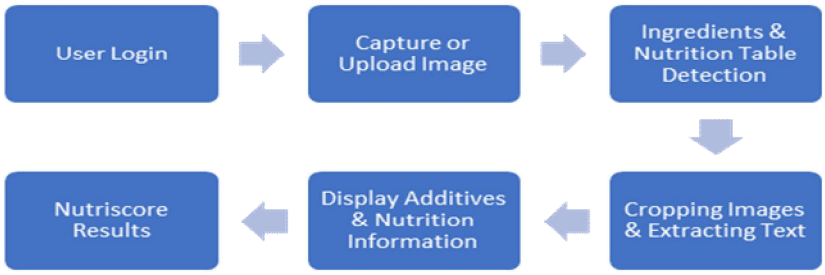



Figure 3. Architecture of the Application

1) *User Login*: The user login process involves the authentication of a user's identity to grant access to the application. This step ensures that only authorized users can utilize the features of the platform, maintaining data security and personalization
2) *Capture or Upload Image*: This step allows the user to either capture a new image using the device's camera or upload an existing image from the local storage. This functionality enables the user to input the image data into the system for further processing.
3) *Ingredients and Nutrition Table Detection*: We use our fine-tuned EfficientDet model for this purpose.
4) *Cropping Images & Extracting Text*: The system automatically crops the relevant sections of the image, focusing on the ingredients and nutrition tables. It then utilizes PaddleOCR [3] to extract text from these cropped sections, converting the textual information into a machine-readable format for further analysis.
5) *Display Additives & Nutrition Information*: After the detection of relevant information, the system displays the identified additives and detailed nutrition information to the user. This step provides valuable insights into the contents of the product, assisting users in making informed decisions about their dietary preferences and requirements.
6) *Nutriscore Results*: Using the extracted nutrition data, the system calculates the Nutriscore, a nutritional rating system, to provide a comprehensive evaluation of the product's overall nutritional value. This step offers users a simplified and standardized way to assess the healthiness of the food product, promoting healthier food choices and fostering a better understanding of nutritional content.

## V. WORKING MODULE

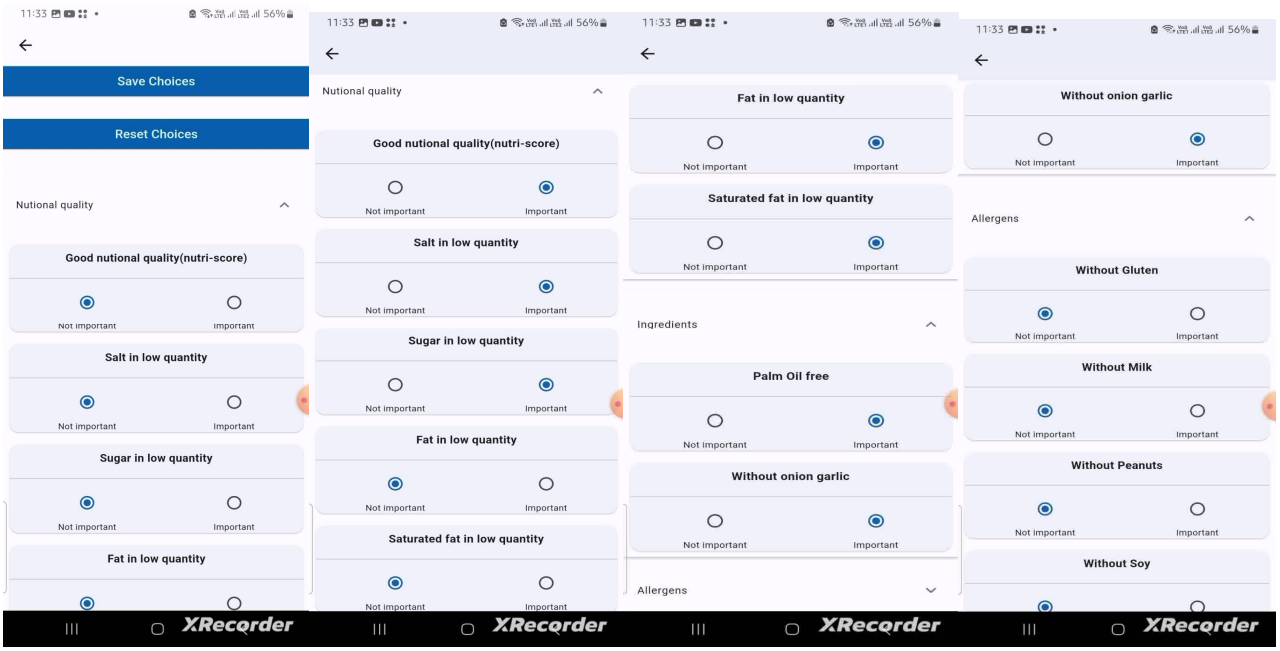



Figure 4. Choose User Preferences

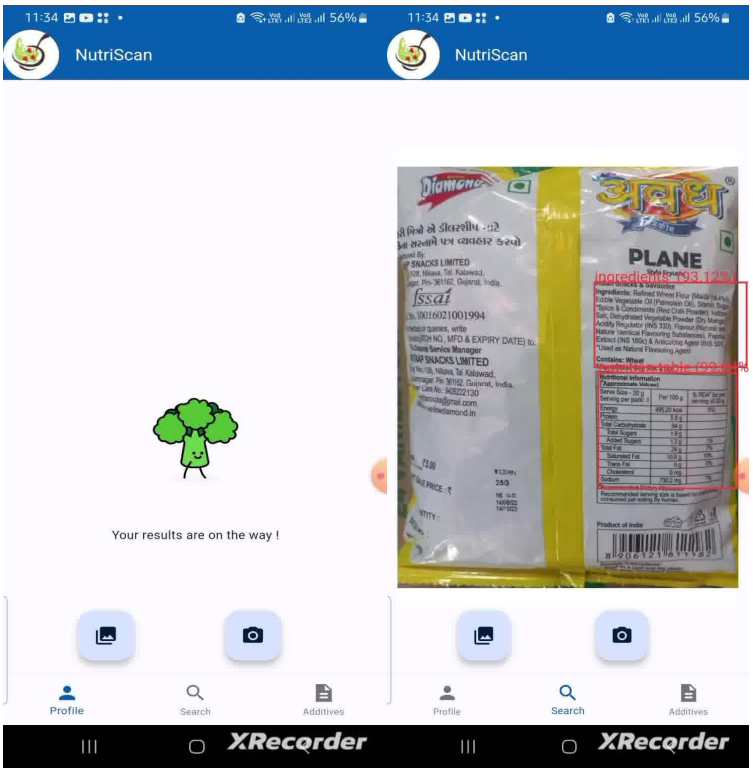
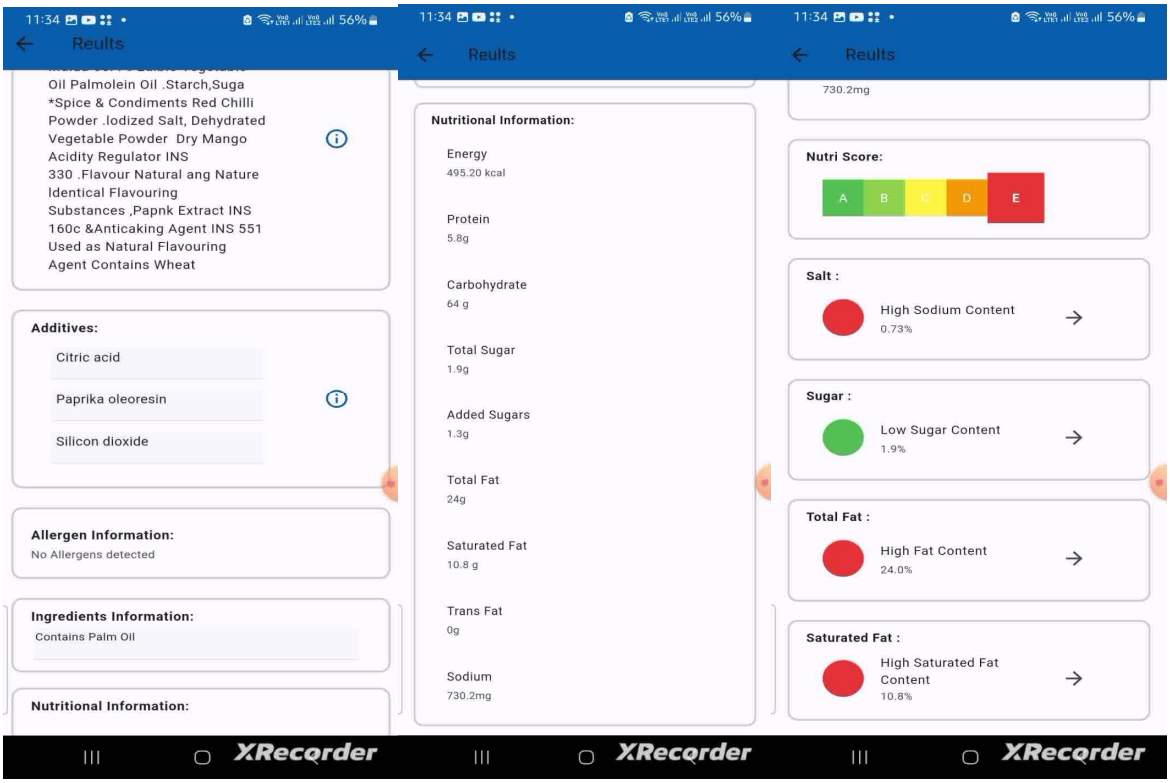



Figure 5. Select the image from gallery or capture it from camera after that the model detects the nutrition table and ingredients.



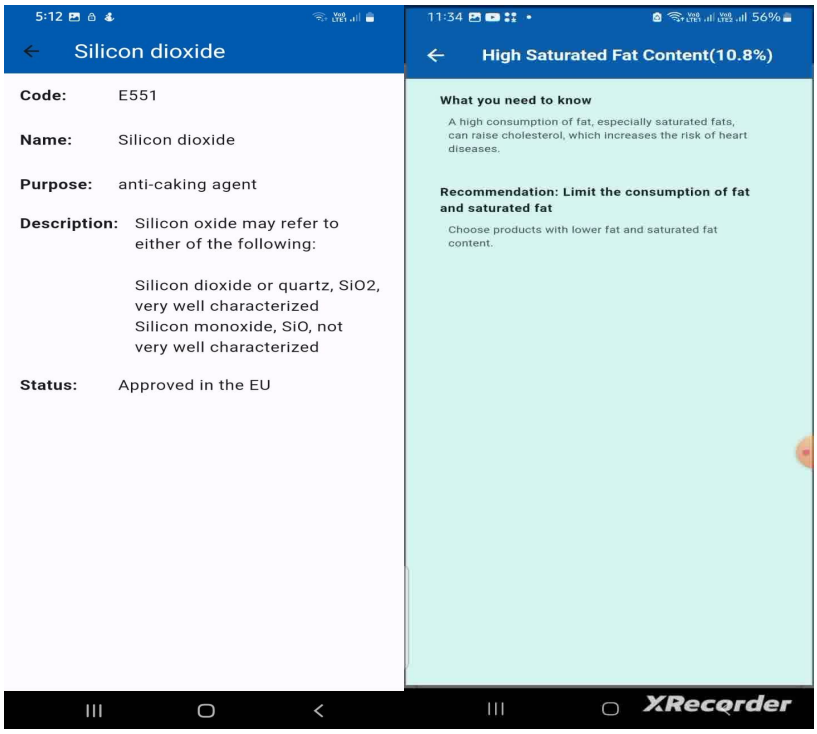



Figure 6-7. This is the result we get after the extraction of information from the cropped images.

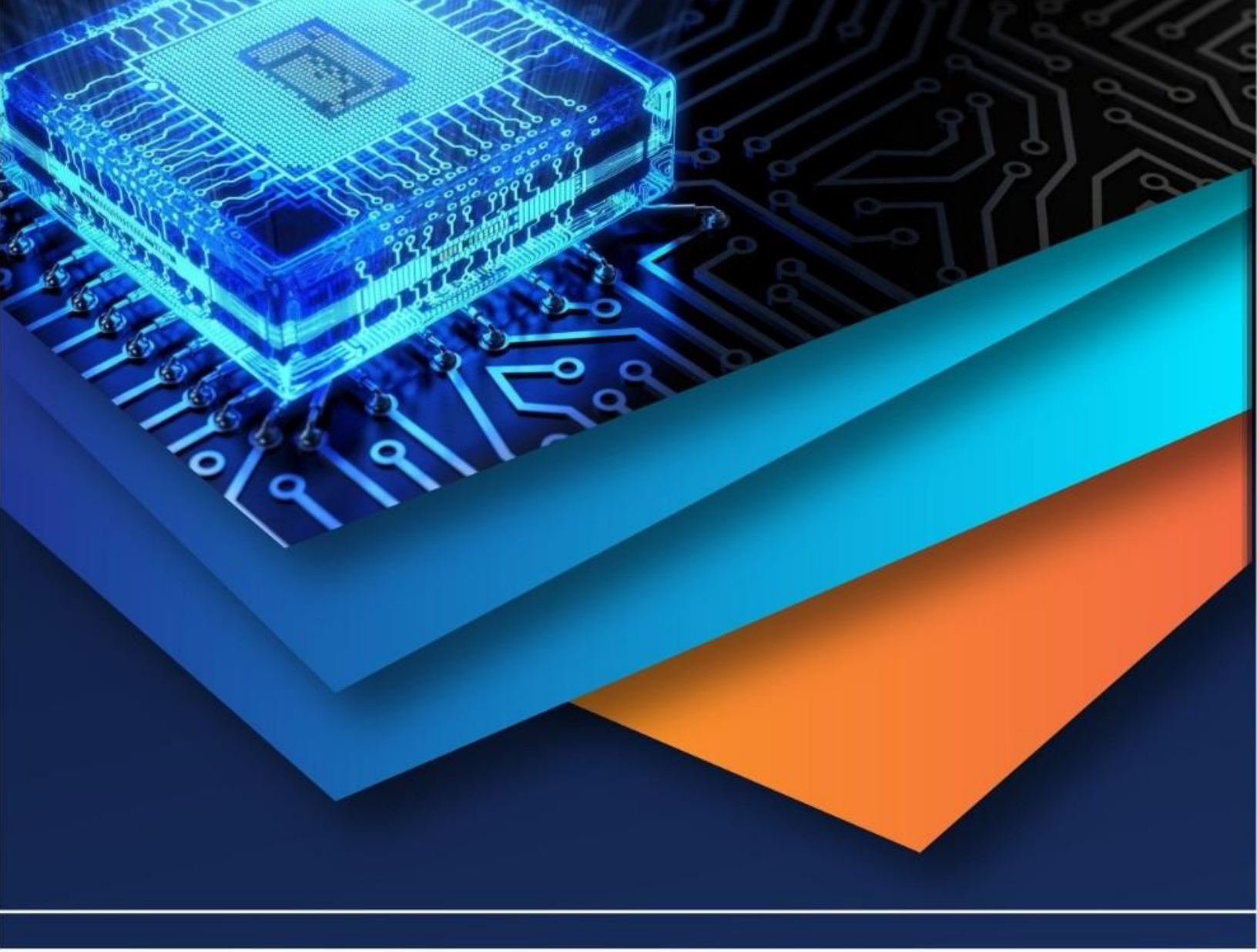
1) *Additives*: The Additives section lists the artificial or natural substances added during processing, helping users understand the ingredients used and make informed choices about the products they consume.
2) *Allergen Information*: This segment outlines potential allergens, such as gluten, dairy, nuts, and soy, enabling individuals with allergies to identify products that may pose health risks.
3) *Ingredients Information*: Providing a comprehensive list of components and their quantities, this section helps users assess the product's quality and nutritional value before consumption.
4) *Nutritional Information*: This section presents vital information about the key nutrients, including carbohydrates, proteins, fats, vitamins, and minerals, allowing users to comprehend the food product's nutritional content at a glance.
5) *Nutriscore*: A color-coded grade from A to E assesses the food product's overall nutritional quality, guiding users towards healthier food choices and promoting balanced diets.

## VI. CONCLUSION

In conclusion, our research represents a significant stride towards bridging the gap between the physical world of packaged food items and the digital realm of automated recognition. With the development of our EfficientDet Customized model, fine-tuned on a diverse and extensive dataset, we have achieved exceptional accuracy and efficiency in nutrition table and ingredient recognition. The integration of food additive information from Wikipedia further enhances the utility of our system, providing users with valuable insights into the products they encounter. Our work underscores the potential for machine learning in enhancing user experiences and promoting informed dietary choices. By offering a seamless and informative bridge

## REFERENCEs

[1] M. Tan, R. Pang and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 10778-10787, doi: 10.1109/CVPR42600.2020.01079.
[2] T. -Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 936-944, doi: 10.1109/CVPR.2017.106.
[3] Tan, Mingxing and Quoc V. Le. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." ArXiv abs/1905.11946 (2019): n. pag.
[4] Li, Chenxia et al. "PP-OCRv3: More Attempts for the Improvement of Ultra Lightweight OCR System." ArXiv abs/2206.03001 (2022): n. pag.
[5] Wikipedia contributors. "Nutri-Score." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 16 Oct. 2023. Web. 1 Nov. 2023.
[6] Wikipedia contributors. "Food additive." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 15 Oct. 2023. Web. 1 Nov. 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)