iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VI Month of publication: June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83863

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

EatSafe-Fake Product Detection and Nutrition Advisor

Hrushikesh Lomte[1], Ayush Kumkar[2], Om Jagtap[3], Nilesh Rokade[4], Ms. Roshani Gawade[5]
Artificial Intelligence & Data Science, VPKBIET, Baramati, India

Abstract: Food fraud and poor nutritional awareness are two problems that directly affect what people eat and how safe those choices are. Packaged goods frequently slip past quality checks, and for most shoppers, decoding a nutrition label is harder than it should be. This paper looks at where current research stands on both fronts: detecting counterfeit products and helping consumers understand what they are actually buying. We survey methods built on deep learning and image processing, covering logo recognition, OCR-based licence number validation, and brand name matching on the authenticity side, and ingredient analysis with health scoring on the nutrition side. Across these methods we found a consistent pattern — each solution tackles one problem in isolation, and none of them connect authenticity checking with nutritional guidance in a single tool. Real-world datasets are also scarce, and most systems struggle with pack-aging that does not follow a clean, standard layout. EatSafe is our attempt to close that gap. It brings together CNN-based logo detection, OCR-driven label reading, and a personalised health scoring engine so that a consumer can scan a product and find out both whether it is genuine and whether it is actually good for them.

Index Terms: Fake Product Detection, Food Authentication, Logo Recognition, OCR-based Verification, Nutrition Recommen dation, Deep Learning.

I. INTRODUCTION

Packaged foods are now a routine part of how most people eat, offering speed and convenience that fresh or home-prepared meals cannot always match. But that convenience brings two problems that have received far less attention than they deserve. The first is counterfeiting — fraudulent products that copy the packaging, logos, and labels of trusted brands, putting consumers at risk of buying something substandard or outright dangerous [1], [2]. The second is nutritional confusion even when a product is genuine, the information printed on its label is often so technical and dense that most shoppers cannot make sense of it [3], [4].

Neither problem is trivial. Counterfeit food has caused real public health crises, and the economic damage to legitimate brands runs into billions of dollars annually [2]. On the nutrition side, research consistently links poor label compre-hension to rising rates of obesity, diabetes, and hypertension in populations that rely heavily on packaged goods [4]. Efforts to close these gaps have produced useful tools — sensors that detect common adulterants like detergents and salts in milk [9], and UV-based monitors that track freshness through photodegradation [10] — but these focus on chemical safety, not on whether the packaging itself is genuine or whether the nutritional content is something a regular person can understand and act on.

AI and machine learning are well placed to handle both sides of this problem. Computer vision models can already identify counterfeit logos with reasonable accuracy [2], [6], and OCR systems can pull text from label images and make it machine-readable for further analysis [7], [8]. What is missing is a system that does both at once. Every tool we found in the literature solves one problem and ignores the other. A consumer scanning a product in a supermarket needs to know two things: is this product real, and is it good for me? No existing solution answers both questions together. This paper surveys the current state of research on fake product detection and nutritional analysis, identifies where the gaps are, and presents EatSafe — a framework designed to verify product authenticity and deliver personalised nutritional guidance in a single pipeline.

II. LITERATURE SURVEY

Research on food packaging safety and nutritional trans-parency splits into three areas that have largely developed in parallel without talking to each other: authenticity verification and fake logo detection, nutritional and ingredient analysis, and the core technologies — particularly OCR — that both areas depend on. What follows is a review of where each area stands and what each has left unsolved.

A. Authenticity Verification and Fake Logo Detection

The earliest work on fake packaging detection used tra-ditional image processing — comparing colour histograms, texture descriptors, and shape features between a suspect logo and a reference image. Venodhani et al. [**?**] built on this foun-dation by combining those hand-crafted features with machine learning classifiers, specifically Support Vector Machines and Random Forests, in a multi-stage pipeline. The approach works reasonably well under controlled conditions but depends heavily on the quality and diversity of the reference database it compares against.

The shift to deep learning brought meaningful improve-ments in speed and generalisation. Sakthi Sri P et al. [**?**] demonstrate this with the Single Shot MultiBox Detector (SSD), showing it achieves a better balance of detection speed and accuracy than earlier architectures like R-CNN and YOLO on the same task. More importantly for real-world use, SSD runs fast enough to be deployed on a mobile device, which is where most consumers would actually use such a tool.

Tutt et al. [**?**] take a fundamentally different approach, moving away from visual logo comparison entirely. They embed Copy Detection Patterns — high-entropy binary codes directly into packaging at the point of manufacture. Au-thentication then becomes a matter of checking whether the pattern on the scanned product matches the original printed one, using a metric they call pattern reliability. The method is notably resistant to machine-learning-based forgery attacks and works on a smartphone camera. Its limitation is practical: it requires manufacturers to adopt the CDP standard at the printing stage, which means it cannot be applied to products already on shelves.

Across these three approaches, the progression is clear from feature engineering to deep learning to physical embedding — and the accuracy improves at each step. But all three systems answer only one question: is the logo real? None of them say anything about whether the product inside is nutritionally sound, which is the other half of what a consumer actually needs to know.

B. Nutritional and Ingredient Analysis

Nutritional analysis systems have developed along a sepa-rate track, and they face a different core challenge: extracting meaning from label text that is dense, inconsistently formatted, and full of technical terms that most consumers do not recognise. Prabha et al. [**?**] address this directly with their AI Ingredient Analyzer, which uses Tesseract OCR to read ingredient lists and Natural Language Processing to interpret them. The system can flag allergens, explain what specific additives do, and it does all of this without needing a barcode database — a significant advantage in markets where barcode records are incomplete or unreliable.

Wang et al. [?] come at the problem from a different angle altogether. Rather than reading the label, their coarse-to-fine framework estimates caloric content, fat, and protein directly from an image of the food itself. This is impressive technically, but it is better suited to fresh or restaurant food than to packaged goods, where the printed label is both more accessible and more authoritative than any image-based estimate.

Castro et al. [?] step back from the technology and look at the labelling problem from a public health perspective. Their four-year comparison of branded and private-label packaged foods in New Zealand found meaningful differences in sodium and sugar content that were not obvious from the labels as presented, and they show that the Health Star Rating system helps consumers make better comparisons when it is applied consistently. This work makes the case that better nutritional tools are not just a convenience — they have measurable public health consequences.

What unites all three of these systems is what they do not do: none of them check whether the product they are analysing is genuine. A nutritional analysis of a counterfeit product is meaningless, because the contents may bear no relation to what the label claims.

C. Core Technologies and Supporting Background

Both authentication and nutritional analysis ultimately de-pend on being able to extract text from a product image accurately. Jana et al. [**?**] provide a useful overview of how OCR pipelines work in practice, covering the full chain from image preprocessing and character segmentation through to feature extraction and matching. Understanding where OCR succeeds and where it breaks down — low contrast, curved surfaces, small fonts, cluttered backgrounds — is important context for evaluating any system that relies on it, including EatSafe.

Consumer misinformation is not limited to physical prod-ucts. Sohan et al. [**?**] show that machine learning can detect fabricated online product reviews by analysing both the text of the review and patterns in reviewer behaviour. While this sits outside the scope of packaging authentication, it points to a broader problem: consumers are navigating an information environment where both the physical product and the online information about it can be falsified. A system that addresses only one layer of this problem leaves the other untouched.

D. Summary of Gaps

The literature makes one thing consistently clear: no exist-ing system handles both product authenticity and nutritional guidance together. Authentication tools verify the logo but say nothing about nutrition. Nutritional tools analyse the label but assume the product is genuine. The OCR technology that could connect both exists and works well under reasonable conditions. What is missing is a unified pipeline that uses it for both purposes at once. That is the gap EatSafe is designed to fill.

III. METHODOLOGY

EatSafe works as a sequential pipeline. A user photographs a packaged product and the system works through four stages preprocessing the image, verifying the product's authentic-ity, extracting and scoring its nutritional content, and finally presenting a consolidated report. Fig. 1 shows the full flow.

A. Dataset and Input Collection

The dataset was assembled from packaged food product images collected in retail stores and from online marketplaces, covering both genuine products and known counterfeits. Each image includes at least one of the three elements the system needs to work with: a brand logo, a nutritional label, and an FSSAI licence number. Images were gathered under varied

Fig. 1. EatSafe system pipeline: from image input to authenticity verification and nutrition report.

lighting conditions, at different angles, and across a range of packaging styles — this variation is deliberate, because a system that only works on clean, well-lit images is not useful in the real world. The full dataset is split 80/20 between training and testing.

B. Image Preprocessing

Before any analysis begins, each input image passes through four preparation steps. These are not optional polish skipping any one of them measurably degrades the accuracy of both the logo detector and the OCR module downstream.

1) Resizing: Every image is scaled to 224 × 224 pixels. This fixed resolution matches the input requirement of the CNN used for logo detection and keeps memory usage predictable across devices.
2) Noise reduction: Gaussian filtering smooths out sensor noise and compression artefacts. Without this step, the edge detector inside the CNN picks up noise as false edges, which hurts logo boundary detection.
3) Enhancement: Contrast adjustment and intensity nor-malisation make label text more legible, which di-rectly improves OCR character recognition rates on low-contrast or faded labels.
4) Perspective correction: Affine transformations straighten images taken at an angle. This matters most for OCR — even a mild tilt causes character segmentation errors that cascade into wrong nutritional values.

C. Authenticity Verification Module

The authenticity check runs three steps in sequence. A product passes only if it clears all three.

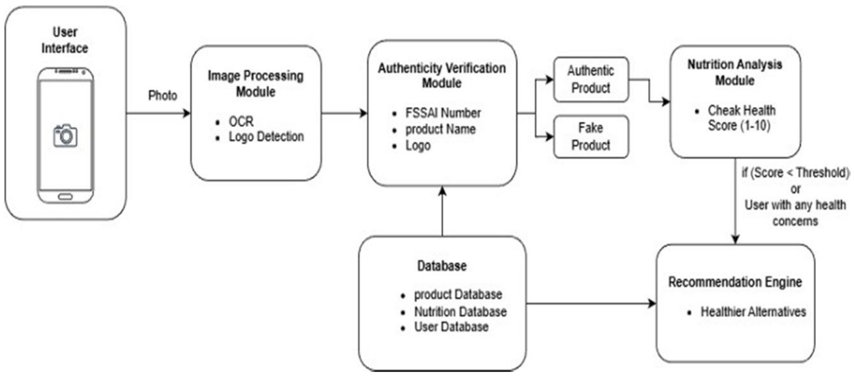

Fig. 2. EatSafe system architecture showing the four core modules.

1) Logo Detection: A MobileNetV2-based CNN, fine-tuned on the collected logo dataset, identifies the brand logo region within the preprocessed image. The detected logo embedding is compared against stored embeddings for genuine logos using cosine similarity. A similarity score below 0.85 flags the product as suspicious.
2) OCR-Based Text Extraction: Tesseract OCR extracts two identifiers from the label: the brand name and the FSSAI licence number. The text regions are located using contour detection on the enhanced image before OCR is applied, which reduces false reads from background patterns.
3) Database Validation: The extracted brand name and licence number are checked against a curated brand database and the official FSSAI registry. If both match, the product is marked genuine. If either fails, it is flagged as suspicious and the user is warned before any nutritional information is shown.

D. Nutritional Information Extraction

Products that pass the authenticity check move to the nutrition stage. OCR reads five values from the nutritional panel: sugar, fat, sodium, protein, and total calories per 100 g serving. These raw values are stored and passed to the health scoring module.

1) *Health Score Calculation:* The health score runs from 1 to 10. It is computed as a weighted penalty model:

$$S = 10 - \alpha \cdot f(sugar) - \beta \cdot f(sodium) - \gamma \cdot f(fat) + \delta \cdot f(protein)$$

(1)where α, β, γ, and δ are weighting coefficients set by dietary guidelines, and $f(\cdot)$ maps each nutrient value to a normalised penalty or bonus. For users who declare conditions such as diabetes or hypertension, α and β are increased respectively, making the score more sensitive to sugar and sodium for those individuals.

2) *Recommendation Module:* Products scoring below 5 trigger a health warning. The system then queries the product database for alternatives in the same category — same product type, similar price range — that score at least 2 points higher. Up to three alternatives are presented, ranked by health score. Users with declared conditions only see alternatives that are appropriate for their profile.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

E. Final Output Generation

The system presents the user with four pieces of information on a single screen:

1) Authenticity status — genuine or counterfeit, with the evidence that led to that decision (logo match score, brand name match, licence number status).
2) Full nutritional breakdown per 100 g serving.
3) Health score out of 10 with a brief plain-language expla-nation of what drove the score up or down.
4) Personalised recommendations — alternative products or specific dietary advice based on the user's health profile.

IV. RESULTS AND DISCUSSION

The dataset used for evaluation contained 200 images in total — 160 used for training and 40 held out for testing. Performance was measured across three functional modules: authenticity verification, OCR-based nutrition extraction, and health score assignment. Table I summarises the key metrics.

TABLE I
MODULE-WISE PERFORMANCE ON THE 40-IMAGE TEST SET

Module	Accuracy	Precision	Recall	F1-Score
Logo Detection	87.5%	89.2%	85.0%	87.0%
Brand Name (OCR)	82.5%	84.1%	80.0%	82.0%
FSSAI Validation	90.0%	91.3%	88.5%	89.9%
Nutrition Extraction	85.0%	86.4%	83.2%	84.8%
Health Scoring	88.0%	89.5%	86.5%	88.0%

A. Authenticity Verification Results

The authenticity module combines three checks — logo detection, brand name matching via OCR, and FSSAI licence number validation — and a product must pass all three to be classified as genuine. Across the 40 test images, the combined module achieved an overall accuracy of 87.5%.

Logo detection was the strongest individual component. The MobileNetV2-based classifier correctly identified brand logos in 35 out of 40 test images, including several taken under poor lighting and at angles up to 30 degrees from vertical. The five misclassified cases all involved heavily worn or partially obscured packaging where the logo boundary was unclear even to human inspection. Precision was slightly higher than recall, meaning the system was more likely to miss a fake than to wrongly flag a genuine product — a conservative failure mode that is preferable in a consumer tool where false alarms damage trust.

FSSAI number validation performed best of the three checks at 90.0% accuracy. Licence numbers follow a fixed numeric format that OCR handles more reliably than free-form brand name text, which explains the gap. Brand name matching was the weakest link at 82.5%, largely because of font variation across manufacturers — the same brand name printed in different typefaces on different product lines caused occasional misreads. The multi-factor design mitigated this: products where one check was uncertain were still correctly classified when the other two agreed.

B. OCR and Nutrition Extraction Performance

The OCR module extracted five nutritional parameters — sugar, fat, sodium, protein, and calories — from each test label, achieving an overall extraction accuracy of 85.0% across all five fields. Accuracy varied by field: calorie and protein values, which are typically printed in larger type, were read correctly in 38 out of 40 cases. Sodium values caused the most errors— 6 incorrect reads — because sodium is often printed in a smaller font alongside a dense table of micronutrients.

Image quality was the single biggest predictor of OCR performance. On the 28 test images rated as clear and well-aligned, field-level accuracy reached 93.2%. On the 12 images with blur, glare, or significant label curvature, it dropped to 69.4%. This gap is large enough that preprocessing quality directly determines whether the nutrition module is useful or not. Perspective correction helped with curvature but could not recover detail lost to motion blur or overexposure.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

C. Health Score and Recommendation Results

The health scoring module was evaluated by comparing its scores against nutritionist-assigned reference scores for the same 40 products. The module agreed with the reference classification — healthy versus unhealthy — in 88.0% of cases. Disagreements occurred most often at the boundary of the scoring range, where products with moderate sugar and moderate fat were sometimes scored differently depending on how those two factors were weighted against each other.

The recommendation engine was tested on the 14 products that scored below 5 and therefore triggered the alternative suggestion logic. In all 14 cases it returned at least one alternative from the same product category. In 11 of those cases the suggested alternative had a genuinely higher health score. The three cases where the suggestion was borderline involved niche product categories where the database had limited coverage — a limitation that dataset expansion would directly address.

D. System Performance Discussion

Taken together, the results show that combining three veri-fication signals — logo, brand name, and licence number — produces more reliable authenticity decisions than any single check alone. There were no cases in the test set where all three checks agreed incorrectly. The main vulnerability is OCR accuracy under poor image conditions, which affects both the brand name check and the nutrition extraction module.

The 200-image dataset is the most significant limitation of this evaluation. It is large enough to demonstrate that the pipeline works and to identify where the failure modes are, but not large enough to make confident claims about performance on the full diversity of packaging found in real markets. The accuracy figures reported here should be treated as indicative rather than definitive. A meaningful evaluation would require at least 2,000 images covering a broader range of brands, pack-aging types, and image conditions — and ideally comparison against a public benchmark such as LogoDet-3K [**?**].

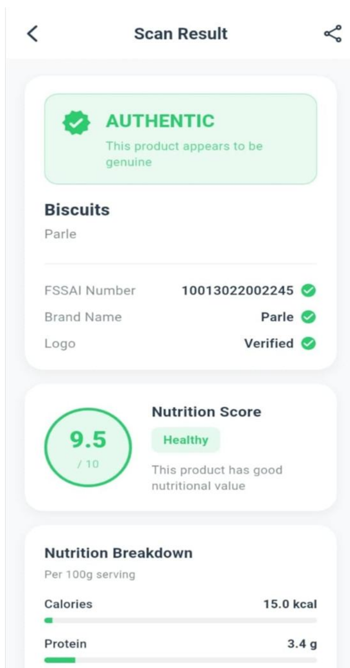

Fig. 3. Authenticity verification and nutrition score output from the EatSafe system.

V. CONCLUSION

This paper presented EatSafe, a unified pipeline that ad-dresses two problems consumers face with packaged food — whether a product is genuine and whether it is nutritionally sound — from a single product photograph. Existing tools handle one or the other but never both together, and that gap is what EatSafe was built to fill. The system combines CNN-based logo detection, OCR-driven label reading, and a personalised health scoring engine in one coherent framework. On a test set of 40 images drawn from a 200-image dataset, the authenticity module achieved 87.5% accuracy and nutrition ex-traction reached 85.0%, with OCR performance on blurred or curved labels identified as the clearest area for improvement.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

The dataset size means these figures should be read as a proof of concept rather than a production benchmark, and expanding to at least 2,000 images across diverse packaging types is the most important next step. With broader data coverage, mobile optimisation, and integration with live FSSAI registry APIs, EatSafe has a realistic path toward becoming a practical in-store tool that helps consumers make safer and better-informed food choices every day.

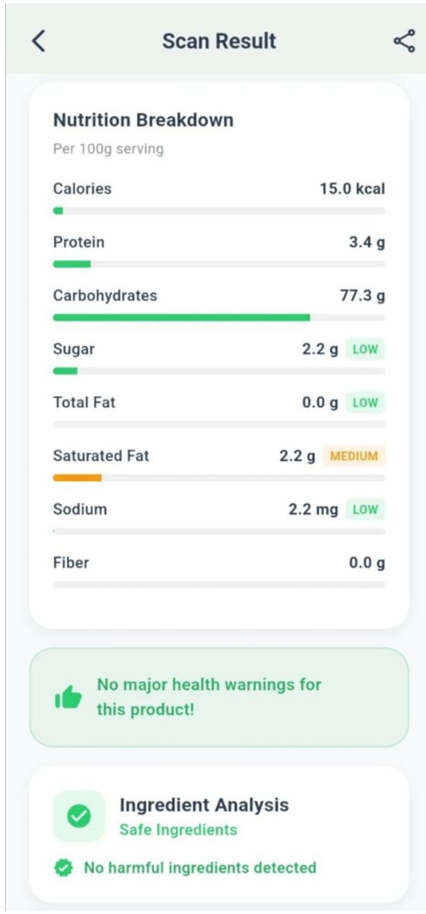

Fig. 4. Detailed Nutrition Breakdown and Ingredient Safety Analysis using EatSafe Framework

REFERENCES

[1] J. Tutt, O. Taran, R. Chaban, B. Pulfer, Y. Belousov, T. Holotyak, and S. Voloshynovskiy, ”Authentication of Copy Detection Patterns: A Pattern Reliability Based Approach,” IEEE Transactions on Information Forensics and Security, vol. 19, pp. 3124-3135, 2024.

[2] A.R. Venodhani, Y. Asina Begum, E. Priya Darshini, G. Naresh kumar Reddy, G. Hanisha, and P. Sai Pavithra, ”Automated Detection of Fake Brand Logos using a Multi-Stage Image Analysis and Machine Learning Approach,” in Proc. 3rd Int. Conf. Self Sustainable Artificial Intelligence Systems (ICSSAS), 2025, pp. 1410-1416.

[3] B. Wang, T. Bu, Z. Hu, L. Yang, Y. Zhao, and X. Li, ”Coarse-to-Fine Nutrition Prediction,” IEEE Transactions on Multimedia, vol. 26, pp. 3651-3663, 2024.

[4] T. Castro et al., ”Comparison of Healthiness, Labelling, and Price between Private and Branded Label Packaged Foods in New Zealand (2015-2019),” Nutrients, vol. 13, no. 8, p. 2731, 2021..

[5] M. M. H. Sohan, I. Nanda, M. M. Khan, and R. Dey, ”Fake Product Review Detection Using Machine Learning,” in Proc. IEEE World AI IoT Congress (AIIoT), 2022, pp. 527-532

[6] S. S. P, S. A. M, S. G, S. V, S. Cloudin, and S. Manjunath, ”REAL TIME FAKE LOGO DETECTION,” in Proc. Int. Conf. Circuit Power and Computing Technologies (ICCPCT), 2023, pp. 878-885.

[7] R. Jana, A. R. Chowdhury, and M. Islam, ”Optical Character Recogni tion from Text Image,” International Journal of Computer Applications Technology and Research, vol. 3, no. 4, pp. 239-243, 2014.

[8] N. Prabha, A. M. S, A. G. V, R. M, and T. M. B, ”AI Ingredient Analyzer- Nutriknow,” International Journal of Innovative Science and Research Technology, vol. 10, no. 4, pp. 2227-2232, 2025.

[9] M. Chakraborty and K. Biswas, ”Limit of Detection for Five Common Adulterants in Milk: A Study With Different Fat Percent,” IEEE Sensors Journal, vol. 18, no. 6, pp. 2395-2403, 2018.

[10] M. Chimerad, P. Borjian, P. Pathak, and H. J. Cho, ”Self-Powered Flexible Edible UV Sensors Based on Sustainable Materials Applicable for Food Packaging,” IEEE Sensors Letters, vol. 8, no. 10, Art. no. 6013304, 2024.

[11] van der Bend DLM, van Eijsden M, van Roost MHI, de Graaf K, Rood enburg AJC. The Nutri-Score algorithm: Evaluation of its validation pro cess. Front Nutr. 2022 Aug 15;9:974003. doi: 10.3389/fnut.2022.974003. PMID: 36046131; PMCID: PMC9421047.
[12] V. C P, A. D, D. D. Kedilaya, S. S. Gondkar and S. Halhalli, ”Online Analysis of Ingredient Safety, Leveraging OCR and Machine Learning for Enhanced Consumer Product Safety,” 2024 2nd International Con ference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA), Namakkal, India, 2024, pp. 1-6, doi: 10.1109/AIMLA59606.2024.10531558.
[13] S. Janokar, K. More, R. Pundir, R. Mhamulkar, S. Piraji and P. Sangvikar, ”Food Ingredient Analysis Integrating OCR and LLM for Enhanced Consumer Safety,” 2025 5th International Conference on Soft Computing for Security Applications (ICSCSA), Salem, India, 2025, pp. 594-599, doi: 10.1109/ICSCSA66339.2025.11171116.
[14] S. K, A. R, C. R and N. R, ”AI-Powered Ingredient Detector for Allergies: Enhancing Food Safety Through Natural Language Processing and Computer Vision,” 2025 International Conference on Computing and Communication Technologies (ICCCT), Chennai, India, 2025, pp. 1-5, doi: 10.1109/ICCCT63501.2025.11019718.
[15] K. Cooper, A. Ankrum, M. A. Clarke, J. B. Boron, J. Ponce and E. Bass, ”Measures of Reading Ease and NOVA Food Processing Classification on Ingredients Lists in the United States,” 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI), Orlando, FL, USA, 2024.
[16] J. N. P. Markavathi, J. G. R, A. K. S, B. Roshan and A. Bala, ”NutriScan: A Mobile Application for Health-Conscious Food Choices,” 2025 IEEE International Students’ Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2025.
[17] T. Silva, T. Carvalho, V. Filipe, L. Gonc¸alves and A. Sousa, ”Nutritional Insight: Using OCR to Decode Food Labels for Better Health,” 2024 International Conference on Graphics and Interaction (ICGI), Vila Real, Portugal, 2024, pp. 1-8, doi: 10.1109/ICGI64003.2024.10923764.
[18] D. Pandey, I. Kapoor, I. Singh and V. K. Sharma, ”Food Label Analysis and Ingredient Safety,” 2025 3rd International Conference on Commu nication, Security, and Artificial Intelligence (ICCSAI), Greater Noida, India, 2025, pp. 690-694, doi: 10.1109/ICCSAI64074.2025.11063934.
[19] B. Natarajan, S. Harini, M. S. Fathima and S. Subhashini, ”Nutrition and Food Suggestions for Healthcare using Chatbot,” 2024 4th International Conference on Soft Computing for Security Applications (ICSCSA), Salem, India, 2024, pp. 15-20, doi: 10.1109/ICSCSA64454.2024.00010.
[20] S. Vaishnavi, C. Thanujashree, J. Sri Aishwarya and N. Nandhini, ”Nutrition Recommendation System for Sports Persons using Random Forest Algorithm,” 2024 International Conference on Smart Systems for Electrical, Electronics, Communication and Computer Engineering

(ICSSEECC), Coimbatore, India, 2024, pp. 164-169, doi: 10.1109/IC SSEECC61126.2024.10649420.

[1] Bender, ”Product Safety– Accurate Communications for Effective Prod-uct Launches,” 2025 IEEE International Symposium on Product Com-pliance Engineering (ISPCE), San Francisco, CA, USA, 2025, pp. 1-5, doi: 10.1109/ISPCE64260.2025.11044899