# Multi-cuisine and Multi-data Screening of Food quality at Service Points using Object Detection capabilities of YOLO NAS

Sai Sathwik Kosuru
*Computer Science and Engineering*
*B V Raju Institute of Technology*
Narsapur, India
21211a05r6@bvrit.ac.in

Adithya Om Sangani
*Computer Science and Engineering*
*B V Raju Institute of Technology*
Narsapur, India
21211a05t0@bvrit.ac.in

Srinivas Pullaboina
*Computer Science and Engineering*
*B V Raju Institute of Technology*
Narsapur, India
21211a05p9@bvrit.ac.in

Sandhya Baindla
*Computer Science and Engineering*
*B V Raju Institute of Technology*
Narsapur, India
22215a0523@bvrit.ac.in

Lanke Pallavi
*Computer Science and Engineering*
*B V Raju Institute of Technology*
Narsapur, India
pallavi503@gmail.com

***Abstract*—The food industry has seen unequivocal negligence, for years, due to vivid commercialization and fierce market competitions occupying agendas, leading to widespread ignominies since food is dependent on agriculture and health collectively. In this context, The major emphasis is on food processing industries but household-level application still needs to be developed. The main development in this work is the optional use of the Internet of Things, as opposed to the existing systems which implemented multiple IoT devices to trace various physical aspects vis-a-vis food quality where there is no economic feasibility. We hereby propose a computer vision-based model that uses the Electromagnetic spectrum to trace the quality of all the elements, including even the obscured in the dish, and that is integrable with various IoT and cloud applications. The model uses YOLO NAS, fine-tuned on four different data sets, leading to the development of state-of-the-art service point-based applications.**

***Index Terms*—Food quality, Sustainability, Deep Learning, YOLO V8, Object Detection**

## I. Introduction

Food is the primary focus in terms of energy considerations and the regular activities of the human body functions. Besides air being the dominant parameter daily, food forms a crucial component for leveraging the major biological activities and metabolic semantics of the human body. It is hereby essential to note the subsequent aspects of the food service and food quality considerations in terms of large-scale services where the subjects of negligence and unforeseen incidents cover the maximum portion of erroneous events. Though these erroneous events seem small or insignificant from the perspectives of discussions and the prospects of plans of these solutions may seem unrealistic and impractical. Food products contain a mixture of organic components responsible for healthy upbringing of the human digestive system, and unwanted ingredients, whether organic or inorganic, negatively impact human health, thereafter affecting all the interleaved systems of the body. Thus, the anticipation of positive dynamics in the food infrastructure is only made true with relevant solutions.

Deep learning models enable computers to examine the discrete pixel structures of the input images, deciding the semantics/segments, and ultimately generate labels/annotations for future reference by the model at the end of the training process. The Requirement for massive data is the central aspect of the initial parts of this solution, which has been a great challenge owing to the lack of appropriate existing systems. YOLO may be the best suited since other networks are complex, and only support image classifications and the major need in this project is to label the objects in the images. Though images may be used at a higher frequency in this paper to explain the models, videos are of primary concern in the final solution proposed. Video feeds may be analyzed as frames, synonymously images in the process. Food items and other impurities are detected in this process. The solution is helpful and novel since it is subject to real-time, and near-practical constraints at the canteen service locations associated with the real-time constraints and flexibility requirements. Concrete methodologies facilitate almost perfect results in this work and this work is part of that. Full achievement of the desired results may take time and need sufficient resources.

## II. Literature Survey

Plantation of Tomatoes in Nigeria require proper post-harvest care to prevent accidental infestations and diseases leading to grain wastage. Though this is an agricultural issue, there has been some relevance in terms of food quality impacting vegetable prices with tomatoes in this context, thereafter causing organizations to cut off individual items as ingredients that affect food quality in total. Some papers in the past suggest using IoT for storage and transportation purposes which may improve the abovementioned conditions. picking, processing, packaging, and palletizing [3]- These are some of the parameters considered in a few applications. Machine learning [9] [13] [14], Deep Learning [7], and Computer Vision [10] techniques are employed in such cases. The involvement of Artificial Intelligence [6] in food markets benefits both commercial and agricultural organizations where they choose certain applications. Moreover, Food Quality and conditions are aspects of value chain that affects both ecology and the economy. Therefore, it is a matter of consumer opinions and perceptions where they focus on the benefits of the items being served. In canteens which are referred to in the current paper, this aspect is of little significance since their choices decide, but the quality of the food served matters [11]. Hence, this aspect has been considered in some or the other manner in the current work. While the existing works contribute little to food quality detection at the service point, which is important, it has been necessitated for us to introduce the current work where we focused on the tail of the above-mentioned value chain.

Biotechnologies [4], Genome analysis, Bioinformatics, etc., have gained sufficient momentum in recent years, wherein they influence the use of synthetic techniques [15] in the food industry [2] to yield optimal results. This may positively impact the sales and the market value of the industry but fails in providing good quality products [12] to the end consumer where the quality is significantly limited by less power [5] of the artificial techniques used. These current research topics are being implemented on an accelerated scale in the following countries- China, USA, Spain, Italy, Brazil, and India [1] [6]. Still, the impact of any discrepancies [7] affects the whole globe. It may be pretty irrelevant to view all these parameters for a canteen-level or service point-level [8] research. Still, all of these constraints have implications in a positive number of current work's topics, thus highlighting the need for an efficient literature survey for any similar solution in the future. A significant proportion of research trends in this decade is reflected in 2020, as indicated by paper publication figures underscoring the evolving trends of research. Therefore, the researchers in the current work have dedicated a substantial time to a literature survey and Problem statement identification purposes, thereby making this research a survey-based research-oriented technical solution holding sustainable scope for decades.

## III. Proposed Work

### A. Description of The Problem Statement

Sustainability has often found issues and challenges in the food sector with significant effect on minority communities and lower-class families having weak access to enough number of food sources. Untidy places and unhygienic environments add as additional challenges to the above problems. Therefore, the current solution works on the screening process of food quality at the Canteen Service Points using Computer vision and IoT. IoT application has not been considered in the current work but may be added to the future scope of this work.

### B. Objectives

- Optimize food service efficiency and ensure good quality of the food served.
- Develop a system to prevent erroneous circumstances by accident or by intention
- Provide sufficient information about the nutritional data vis-a-vis food service qualifications of the concerned organization.

### C. Working

IoT solutions are often preferred in the majority of the case in food-tech solutions to measure some physical parameters that are left undetected by the software algorithms. These solutions may be feasible in places where there is a liberal disbursal of funds, but this may not be the case everywhere. Deep Learning algorithms may seem economically fit in all possible scenarios, and this is the main proposal from the current solution. Food images are pretty dense, and they obscure many anomalies within, thus hindering the algorithms to detect the anomalies perfectly, which is again an issue. Vibrators may be arranged in such cases but may not be feasible in every organization. Hence, we propose the use of microwave and thermal rays to generate visual versions of the video feeds. This enhances the visual coverage of all the ingredients in the placed food item, thereby exposing even the obscured items. These rays possess high frequencies and use extra-visual information to derive insights into the obtained data. This provided an extended version of visibility, thereby spanning the global scope of the current solution's applications.

Considering the overlapping objects may not only solve the issue since there are many permutations of the input data, and various constraints like lighting conditions, angle of view in which the video/photo is captured, etc. making inputs unpredictable. In this context, Various possibilities are to be considered, and the same was considered in the initial stages of the current research. Adding to this, positional and orientational coordinates, together with spatial dimensional constraints, may deviate the pre-trained learnings of the model from the ideal condition. Computational considerations like that of hardware may slightly limit the efficacy of the system, but these issues may be solved with proper consideration of all the possible scenarios of the input data captured. Therefore, it has been proposed to create a computer vision model, capable of detecting food items, analyzing the input feed, and detecting all kinds of anomalies based on the learnings from the pre-trained models developed as a part of this solution and from the external resources in any kind of situation regardless the environmental constraints posed on the input feed.

### D. Methods

*1)* ***Data set:*** Food data is diversely arranged and covers a variety varieties, nationalities, and cuisines. Therefore, it is not

| | | |
|---|---|---|
| ImageNet | MS COCO | Roboflow |
| Food 11 | Food 5K | Github Repositories |
| Manual | Fine-tuning | |

TABLE I
DATA SETS USED IN THE STUDY

easy to generalize these types of data sets on different models. Hence, we propose a multi-modal and multi-class detection algorithms to be implemented in order to cover the diverse areas. It is futile to involve all the data sets and consider a significant number of classes for detection. Though offensive, this is true since complex data sets easily confuse a model. Hence, we considered only Indian dishes in table I for the current solution, and this is research on a particular area concerned with food delivery and quality.

*2) **Technical Considerations of the Methodology**:* Classes and labels are both distinct but are collocating to a single image. Multi-class and multi-label analyses lead to different semantics in terms of clustered groups. Classifying a whole image may be termed multi-label classification. At the same time, classes may be variant in terms of pixel semantics within the images, which may be alternated as multi-class classification. Object detection may be specified in the latter case which is the main focus of the current solution.

In the current scenario, the detection task may be segmented as two Phases and seen in Figure 1: Food Service points are crowded locations with unspecified constraints and unknown boundaries. Vast region of consideration neutralizes the efficacy of semantic localization in the image, thus affecting the accuracy of the detected objects. Moreover, dimensionality must be reduced in order to save memory, and improve performance. Contour or color-based segmentation may be taken as metrics to classify layers of the food. Regions of interest have to be deeply investigated for various ingredients, and things present in the dish, where anomalies can be sought out. Spatial regions of interest can also be considered since the current scenario is not orientation-specific and there are many things to be considered. Once the region of interest is picked, we need to move to Phase 2-Object Detection and Quality Analysis.

The region of interest is deeply analyzed for objects in table II and all the anomalies are sought out where if that anomalies are serious as per the pre-defined conditions, the user will be alerted immediately thereby preventing accidental ingestion of harmful materials through food. Clear insights like Nutritional information with some additional features describing the detailed health benefits may be implemented for interdisciplinary use cases involving health sciences. Although this is of little importance concerning for public views, introspection of these details may preserve safety measures and hygiene practices taken by the organization. Attributes of the dish may define some associated short- and long-term health effects.

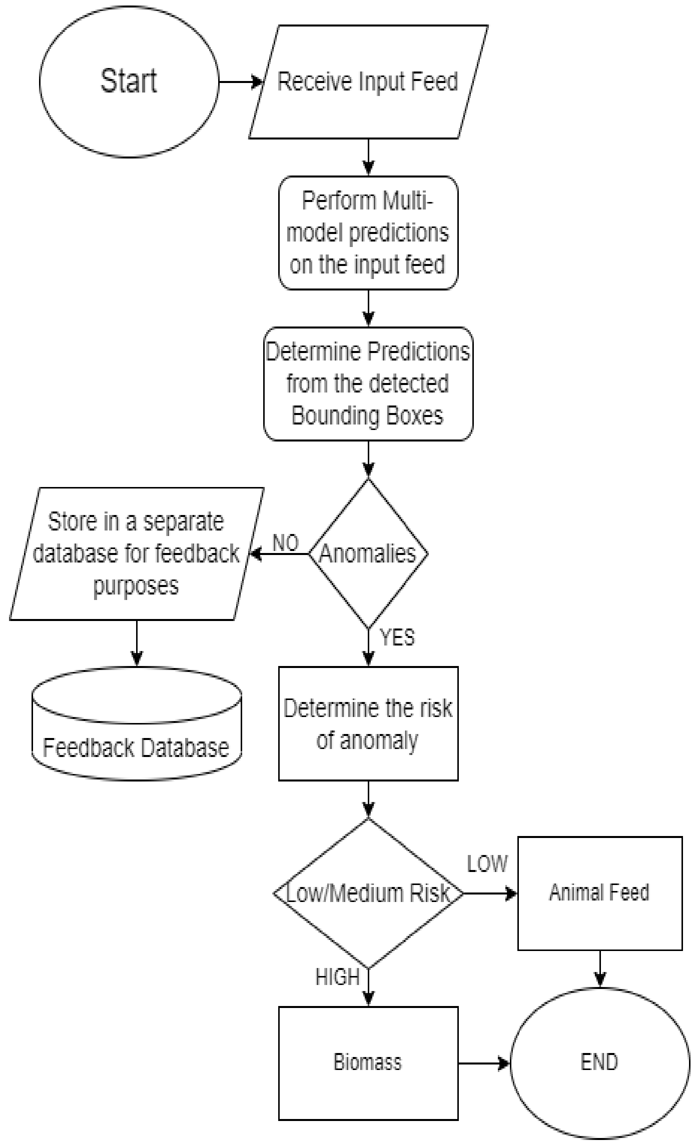


Fig. 1. Process flow of the working system

| | |
|---|---|
| ROI | 60%-80% |
| Items | Food, Dirt, Insects, Frogs, Lizards, Hair |
| Segmentations | Overcooked, Uncooked, Not well cooked |
| Database | Nutritional Information |

TABLE II
TECHNICAL CONSIDERATIONS OF THE WORKING MODEL

*3) Proposed Architecture or model-Working of YOLO NAS:* Image segmentation enables contour-based gradient examinations and color-based semantic differentiations, thereby annotating the distinct overlapping components of the food dish. This may be detected using the suitable features of YOLO NAS pre-trained model. CNN is used for image classification, and Mask RCNN allows us to mask regions with layers of

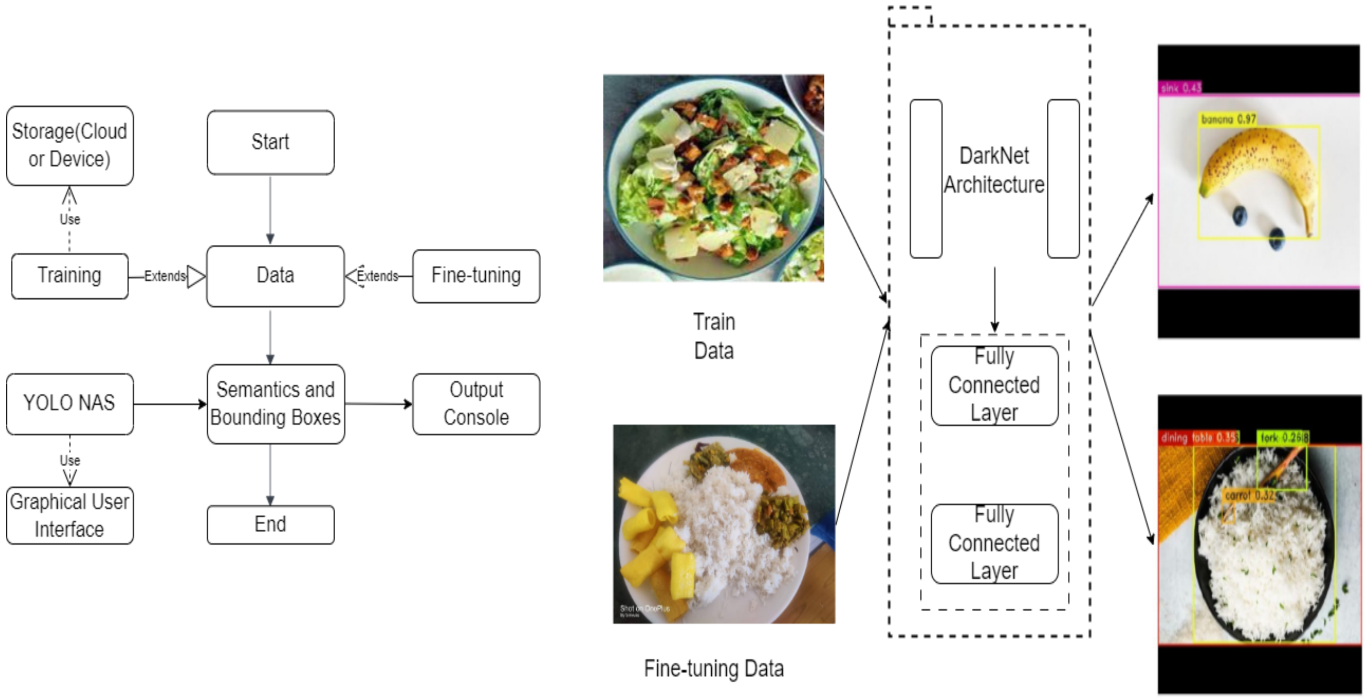


Fig. 2. Hardware Information(Left) and Abstraction Layers(Right) of the Proposed System

learning, thereafter seeking sufficient parameters to detect anomalies. This is a problem of object detection where classification on the whole is unimportant. Hence, CNNs and other deep learning frameworks like ResNet, MobileNet, RNN, etc., may be partially unsuitable to the current solution's domain, underscoring the relevance of YOLO NAS. Henceforth, object detection models may be used for precise definition of this work's purpose. While YOLO has been dominant in this field, there are various other models and frameworks to be considered, but our findings suggest that YOLO is the best of all. YOLO NAS has been the latest in this case, where it implements Neural Architectural Search to find the suitable algorithm. The DarkNet component of YOLO NAS(in figure 2) has been proposed as a standard algorithm in the current work. Therefore, they were consulted frequently throughout the process for various data sets and pre-trained models.

## IV. Result Analysis and Discussions

### A. Normal Images

Latent representations even got better results with YOLO NAS than the previous versions. Normal images without any filters can easily be analyzed for objects in them using this technique. YOLO V8 and YOLO NAS fused are proving effective against inputs that are in proportion with the training images. Note that there is flexibility in terms of orientation and luminous variations in the images. Blurry images may be detected with decimal accuracies, which are still significant. Standard image classification has been subjected to rational changes compared to the existing systems. Existing systems focused on qualifying food and using IoT devices to assess the quality of food, whereas the current solution proposed using vision techniques to determine the food quality.

### B. Thermal Images

This has something that is a novel contribution to the current research. Organism detection, detection in the dark, satellite image classifications, Military use cases, etc., often consider this technology. Still, in terms of food quality, normal images in themselves do not have enough solutions to be contributed to society. Having mentioned that, it is essential to know the importance of Thermal photography to detect thermal aspects of food to assess its quality. IoT applications are often considered to assess the quality, while that may not be feasible in every case. That type of solution will, in turn, complicate the use cases, thus destroying their purpose.

$$Q = mc\Delta T \tag{1}$$

Thermal images are generated considering heat distributions and gradients among the bodies which is examined using the equation 1. Food items contain various substances wherein each body possesses different thermal properties owing to the specific heat capacities(Table III) of the respective substances leading to the convectional, and conduction heat currents between the body, and the surrounding environments. Generally, food items are cooked, or roasted or are frozen for taste and digestion purposes. Hence, there will be significant temperature variations among the ingredients. Specific Heat Capacity is the deciding factor here and since infrared radiations are specifically heat-sensitive. Existing systems have largely been

| Item | Specific Heat Capacity |
|---|---|
| Rice | 0.37 |
| Tomatoes | 0.95 |
| potato | 0.82 |
| Onions | 0.9 |

TABLE III
THERMAL CONSIDERATIONS FOR PREDICTIONS ON THERMAL IMAGES

limited to agricultural sectors in this case where grain, crop, and stored food materials are only inspected and the locations of such activities is centred to industrial complexes, thereby failing in household applicability of the existent solutions. While this may seem insignificant in normal houses, but has a negative impact on the service locations. A future scope of this research can be discussed here, which is to entail using small-scale thermographic cameras on an economic scale.

### C. Radiographic Images

While Electromagnetic spectrum seems complex in this research, their spectacular properties can help researchers in leveraging the effectiveness of their solutions. Radiographic techniques exploit the scattering and reflectivity of the rays to detect obstacles in their paths. This is the same phenomenon used in Radars and same principle employed in Sonars. This is complex for the current work wherein the region of transmission is very small as opposed to large areas of coverage considering interferences, and bandwidths as in communication systems. Food items have a lot of stuff as their ingredients and in some cases, more than a single item may have same heat capacity, thereby faltering the thermographic approach. Reflectivity and scattering principles depend on the size of particles hit by these rays. It is dependent on the relation between the size of the physical particle, and the wavelength of these rays. Therefore, these rays can generate images showing diverse ingredients within the food item under consideration. Radiographic moots may be available, but the presence of mobile solutions is still sparse. Though the images are latent, and largely bichromatic, there is still some feasibility of consideration for this research work. They highlight sized-based metrics of the detected bodies hence showing independence from the other attributes, and knowing the features of all the ingredients will help the models in detecting the foreign bodies because of size variations.

### D. Further Discussions and Impact

Following up from the above result analyses(summarized in table IV), we can state that independence of the above types should never be considered, and the integration of the three types facilitates proper examination of the food quality. Integration of three different devices is still a challenge and a subject of further discussion. This part of discussions is subject to proper coordination from different wings of technology, considering the interleaving of multiple disciplines. It is essential to note the novelties of these work far from the routine landscape of existing works in the food technology sector:

| | |
|---|---|
| Cooked Food | Plates and tables |
| Utensils | Spoons and Fork |
| Insects | Frogs and Lizards |
| Deep fried stuff | Overcooked Materials |

TABLE IV
RESULTS AND POTENTIAL PREDICTIONS

- **Food Quality in terms of visual insights:**Existing systems depend on hardware, i.e., IoT devices, for measuring various physical properties to determine the quality which has no universal feasibility. Therefore, the current system utilizes Computer Vision techniques coupled with EMR concepts to derive valuable insights.
- **Consideration of all associations:**The current system utilizes all the concepts to delve into deeper analytics. Examples include EMR spectra, Non-Cyber Physical technologies, etc.
- **Small-Scale Solutions**Most of the existing systems rely on large-scale applications, but smaller sizes should also be considered. The current solution is a perfect fit in that context.

*1) Description of Results:* Object detection is particularly critical in the current model owing to various overlappings due to the mixture of various ingredients. Simulation-annealing-like behaviour should also be considered pertaining to different cooking methods associated with various recipes, thereby calling for more generalized approaches. Generalized approaches often require multi-dimensional and multi-spanning data points for consideration and training purposes.

**Relevance of YOLO NAS:**
YOLO NAS is a pre-trained model that doesn't require training and hence fine-tuning is performed in our work using locally captured images. Compared to the previous versions of YOLO, this model has more accuracy and is able to identify a vivid range of objects using the MS COCO(Common Objects in Context) data set. It uses DarkNet architecture that assists in properly learning from the training data and obtaining better accuracies. Additionally, NAS stands for Neural Architecture Search which searches for the best architecture for the given problem, hence causing better accuracy.

Limitations of the current system include the following points:

- Though being the latest and advanced, YOLO NAS is quite specific to the clustering and similarity indices of the input pixels. Chromatic aberrations and irregularities may occur while cooking which will negatively impact the current model.
- The model is quite stringent in terms of shape recognition. The Problem arises when two items of similar shapes are placed together, thus confusing the model.

- Thermal and Size properties matching may occur, thereby leading to significant redundancies in predictions.

The above limitations call for further research but thanks to the recent rapid advancements in this field, the researchers anticipate relatively faster results in these aspects in the near future.

*2) Impact and Contributions:* The impact may be felt in terms of further developments in the sustainability sector, and these developments may be both positive and negative. According to research, food coverage across the globe is significantly decreasing daily at an alarming rate. This leads to **Poverty through Food** situations, ultimately making food more valuable than money. Food, Fossil Fuels, Money, etc., get exhausted rapidly. Therefore, they should be used carefully. Food wastage are the primary issues that can be prevented through this system as it detects only anomalies, and the whole food need not be discarded. According to the flow proposed as part of this research, low- or medium-risk anomalies may be utilized for the animal feed, whereas the high-risk for Biogas. The alternative options proposed are again renewable resources, thereby continuing the positive chain. Therefore, this solution is valuable in terms of economy, feasibility, ecology, and in sustainability.

## V. Conclusions and Future Scope

Food quality is a subject of little discussion in the field of technology. This does not provide any overall illustration but deals with a specific part of a colossal process. Agricultural, Horticultural, and Food Processing industries generally rely on these technologies as their primary dependency. Here, it should be noted that small-scale industries and households and other food servicing points should also be considered since they may have a significant effect on the reputation and the businesses of organizations. Hence, this work focuses on providing organisations with sufficient technical support to solve the sustainability issues of food quality and hygiene economically using technology. This research also leaves a positive connotation on the health sector due to intricate interdependencies between lifestyles and various societal sectors. This paper is a proclamation of the development of similar systems in large-scale small-tier industries to improve sustainability aspects in a stochastic manner. While this research covered all the possible aspects, there may arise more issues and areas in the future, that demand the use of sophisticated technologies in the future thereby causing a long-term extension to the current work in the distant future.

As concluded, it is implied that the current work is subject to numerous modifications and updates. It doesn't necessarily mean that the current work becomes obsolete. That's never an issue since this involves the sustainability principles, and all the solutions are accordingly laid as part of the current work. Work may be continued to integrate mobile devices with radiographic moots and thermal sensors to improve scalability exponentially and consistently. We can implement modern frameworks and modules to derive insights with high accuracy. Since the data related to this solution is significantly less, the use of a few data size optimization techniques like data augmentation, Synthetic data generation,and Generative AI simulations may form a part of future work. In other words, there is a possibility to create an entirely new module called "Food Neural Network" for deep learning applications on food.

## References


[1] Grunert, K. G. Food quality and safety: consumer perception and demand. European review of agricultural economics, 32(3), 369-391 2005 https://doi.org/10.1093/eurrag/jbi011

[2] Ellis, D. I., Muhamadali, H., Haughey, S. A., Elliott, C. T., Goodacre, R. (2015). Point-and-shoot: rapid quantitative detection methods for on-site food fraud analysis – moving out of the laboratory and into the food supply chain. Analytical Methods, 7(22), 9401–9414. https://doi.org/10.1039/c5ay02048d

[3] Kamilaris, A., Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

[4] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

[5] Lam, M., Nguyen, T. T. Deep Learning-Based Food Quality Estimation Using Radio Frequency-Powered Sensor Mote. IEEE Access, 8, 88360–88371. https://doi.org/10.1109/access.2020.2993053 2020

[6] Sharma, S., Bisoyee, P., Jathar, J. Food Quality Assurance using Artificial Intelligence: A Review Paper. 2021

[7] Salim, N. O. M., Zeebaree, S. R. M., Sadeeq, M. a. M., Radie, A. H., Shukur, H. M., Rashid, Z. N. Study for Food Recognition System Using Deep Learning. Journal of Physics, 1963(1), 012014. https://doi.org/10.1088/1742-6596/1963/1/012014 2021

[8] Chen, H., Zhang, L., Hu, Y., Zhou, C., Lan, W., Fu, H., She, Y. (2021). Nanomaterials as optical sensors for application in rapid detection of food contaminants, quality and authenticity. Sensors and Actuators B-chemical, 329, 129135. https://doi.org/10.1016/j.snb.2020.129135

[9] Sharif, A., Abbasi, Q. H., Arshad, K., Ansari, S., Ali, M. Z., Kaur, J., Abbas, H., Imran, M. A. (2021). Machine learning enabled food contamination detection using RFID and internet of things system. Journal of Sensor and Actuator Networks, 10(4), 63. https://doi.org/10.3390/jsan10040063

[10] Zhu, L., Spachos, P., Pensini, E., Plataniotis, K. N. (2021). Deep learning and machine vision for food processing: A survey. Current Research in Food Science, 4, 233–249. https://doi.org/10.1016/j.crfs.2021.03.009

[11] Yuan, B. Z., Sun, J. Trend and status of Food Science and Technology category based on the Essential Science Indicators during 2011–2021. Food Science and Technology, 42. 2022 https://doi.org/10.1590/fst.91321

[12] Mohanty, S. P., Singhal, G., Scuccimarra, E. A., Kebaili, D., Héritier, H., Boulanger, V., Salathé, M. The Food Recognition Benchmark: Using Deep Learning to Recognize Food in Images. Frontiers in Nutrition, 9. https://doi.org/10.3389/fnut.2022.875143 2022

[13] Hemamalini, V., Subbaraj, R., Nachiyappan, S., Sambath, M., Devi, T., Singh, B. K., Raghuvanshi, A. (2022). Food quality inspection and grading using efficient image segmentation and Machine Learning-Based system. Journal of Food Quality, 2022, 1–6. https://doi.org/10.1155/2022/5262294

[14] Han, J., Li, T., He, Y., Gao, Q. (2022). Using machine learning approaches for food quality detection. Mathematical Problems in Engineering, 2022, 1–9. https://doi.org/10.1155/2022/6852022

[15] Arashiro, L. T. (2023). Recovery of phycobiliproteins and biogas from microalgae treating wastewater. https://doi.org/10.5821/dissertation-2117-345648