RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt2666

IoT Assisted Multimodal Deep Learning Framework for Real-Time Onion Spoilage Detection

K. Pooja [a, *], A. Reni [a]

[a] Department of Food Processing and Preservation Technology, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore-641043, India

* Corresponding Author Email: pooja_fppt@avinuty.ac.in

DOI: https://doi.org/10.54392/irjmt2666

Received: 13-04-2026; Revised: 22-09-2026; Accepted: 08-10-2026; Published: 30-10-2026

Abstract: Onion storage suffers from major challenges due to post-harvest losses (PHLs) in agricultural supply chains, especially in developing countries where continuous monitoring of surrounding conditions is limited. Standard storage monitoring systems are typically based on traditional static threshold rules (e.g., an excess of 60% relative humidity), and as such, do not have intelligent capabilities to make decisions. To assist with this gap in the current available options for storage monitoring, this study presents a multimodal Internet of Things (IoT) system combined with deep learning for real-time detection of onion spoilage. This unique approach uses environmental data from Internet of Things (IoT) sensors connected via an ESP32 microcontroller to measure temperature, humidity, moisture content, and gas concentration. It also takes images of the onions and uses Convolutional Neural Networks (CNNs) to classify the photos of the onions as they spoil. The application of Grad-CAM visualization of the features contributing to image classification will also assist with the interpretability of the developed models. The evaluation of the proposed framework employed a dataset of 2400 onion images categorized into four classifications: healthy, sprouted, spoiled and with black mold. Precision, recall and F1-score were 91.8%, 93.2% and 92.4%, respectively, for each of the four classifications using the proposed multimodal IoT and CNN framework. Overall, the system achieved an end-to-end latency of approximately 1.3 seconds; therefore, providing almost real-time monitoring of stored onions.

Keywords: Onion, DHT11, MQ3, MQ4, ESP32 microcontroller, Grad-CAM, CNN.

1. Introduction

A significant problem in worldwide agricultural product supply chains is the post-harvest loss of agricultural produce. Onions are especially susceptible to spoilage after harvest because of their high use in food preparation and because microorganisms infect them; temperature fluctuations and moisture changes throughout transit can also affect the quality. An effective grading and storage system for onions, as with most agricultural products grown in large volumes, is essential for successful farming in any country where onions are produced, including India. Currently, traditional sorting and monitoring methods are poor in performance and fail to consistently provide an accurate measure of the success or lack thereof, contributing to lost harvests for those farmers after harvest. However, the Internet of Things (IoT) technology provides opportunities to develop new methods of doing intelligent monitoring of onions and managing their storage [1]. In global architectural market, India ranks as the second largest producer of onion following China. It is estimated that India contributes 26 to 27 million metric tons of onion annually [2]. This discrepancy is mostly caused by issues like pest infestations, plant diseases, environmental stress, and shifting weather patterns. Onions are frequently impacted by pests and diseases such as anthracnose, bulb rot, and seedling damping-off. Each of these issues requires particular management and maintenance practices for effective control and long-term crop health [3, 4].

With the use of fertilizers and pesticides for increasing the crop yield, environment and human health are adversely affected, though they are used to prevent pests and diseases. These pollute the soil, water, harm the ecosystems and leaves the residues on the food consumed. Due to industrial revolution and development in digital technology, these problems are effectively addressed and hence promoted sustainable farming [5]. IoT based technologies offer precise, data driven solutions for the early disease detection in crops, and they provide targeted intervention of the problems. These techniques don't rely on the use of chemicals;

hence they are the best solution in protecting the environment, offering enhanced food safety and increase in the agricultural productivity [6].

IoT sensors are used to continuously monitor the environmental parameters during storage, cultivation, and transportation, while blockchain technology confirms data integrity, and transparency [3, 6]. Together, the mentioned advancements support early spoilage detection, preservation of quality and reduction in risk contamination [7, 8].

Among the various post-harvest illness, black mold disease is caused by Aspergillus niger, a type of fungus that majorly contributes to onion spoilage. Most pathogens proliferate under relatively warm and humid environmental conditions, such as those generally found in either storage areas or places without adequate ventilation [9, 10]. Damage that has occurred to a large quantity of onions may not be obvious until the time comes to sell them; this allows for the continued spread of that damage from the surface of the onion to its bulb portion [11]. DL-based image classification methods have been shown to be effective for early-stage diagnosis of plant diseases [12, 13].

In addition, IoT applications have shown efficiency across a variety of fields and industries. Using this technology, intelligent food packaging and storage conditions have been developed [9]. With the growing adoption of diet management applications, accurate and interpretable food image classification especially for diverse Indian cuisines, has become increasingly important [14]. By integrating DL, reliable decision making across agricultural disease diagnosis and food recognition systems is enabled, enhancing transparency, user-trust and real-world applicability [15].

While researches have been done to support detecting crop diseases with the help of deep learning and IoT-based monitoring systems; there are only limited current monitoring researches concerning detecting crop rot through both environmental based signals and images (non-environmentally based) [16]. Environmentally based solutions utilize measurements to show the trigger points of rot development, but they do not consider any visual clues associated with the stage that caused the visual deterioration. Consequently, research concerning environmental sensor integration and visual inspection for the purpose of detecting onion spoilage is very limited.

Machine learning, deep learning, computer vision, IoT and predictive analysis all allow food manufacturers to monitor how foods are stored, how long they will last and how the product is packaged, in real time. By using integrated sensors, RFID tags and cloud-based monitoring devices, food producers can identify spoilage, contamination, and packaging defects at the time of shipment to improve food safety, traceability and reduce waste of food [17].

The technical innovation of this research project is a trust-weighted multi-modal fusion framework for utilization of environmental sensor readings and CNN-based visual analyses to detect Onion spoilage. In contrast to previous research which has either seen the use of environmental sensors or have utilized image classification independently from one another; therefore, the proposed framework utilizes both sources of complementary data coupled with Grad-CAM explainability and low cost IoT deployment, thus allowing for reliable real time observation of Onion spoilage accurately within practical storage environments.

The balance of this paper will be distributed as follows: In Section 2, prior works regarding IoT-based agricultural monitoring and deep-learning based disease detection will be outlined; In Section 3, a new framework entitled 'Multimodal IoT-CNN' will be presented; Section 4 will provide a description of the experimental setup and evaluation metrics used; Section 5 will provide a summary of the experimental results and a discussion of these results; and finally Section 6 will conclude the paper with an outline of future directions for research in this area.

1.1 Contributions and Objectives of the work

The research presents a confidence-weighted multi-modal fusion framework to provide an enhanced means of fusing environmental sensor data and image predictions generated from convolutional neural networks for the purpose of improving robustness and reliability of onion spoilage detection while allowing for real-time deployment of the IoT and for explainable decisions.

The main contributions made by this research can be summarized as follows. An onion disease diagnosis system is developed based on sensor technology by utilizing various crop-related and environmental variables that include humidity, temperature, toxins, and color changes. There are several sensors used in the system to monitor the health of the crops by using data related to humidity, temperature, gas, moisture, and color. Also, a CNN-based technique is used in the system to classify the onion into four categories: healthy, sprouted, spoiled, and black mold. Additionally, a mobile application is included in the work for ease of access to monitoring and classification results. A Grad-CAM-based framework is adopted in the system to enhance the interpretability of the classification process and allow for timely actions to be taken.

2. Survey

Through the use of IoT and AI technology, traditional warehouses are being transformed into intelligent storage systems that reduce the amount of post-harvest onion waste. By combining gas sensors

with image processing algorithms to detect onion sprouting in real time, a real-time monitoring system with an overall detection effectiveness of 87% can be developed. Using a data-oriented approach to making decisions about how to look after your farm (and how to manage your farm), instead of relying on human judgement using your senses, will allow you to get more accurate decisions, and make it easier to manage your farm with real knowledge. Monitoring the health of your farm and providing useful results, will encourage the implementation of methods to achieve these results. Although it is necessary to conduct studies with gas sensing techniques in a variety of environmental conditions, the proposed method has numerous possibilities for extending, and for implementation in real life in smart agriculture & farming [1].

This research integrates both image processing and an IoT-based monitoring framework to provide a unified system for the early detection of the degeneration of onions. Existing IoT solutions exist to help mitigate these issues but are not well suited to cope with environmental disruptions or measurement error, especially when being compared to existing methods of storage. CNN is utilized via Google Colab to identify onion spoilage with an average accuracy of 87%. In addition, the use of IoT hardware components such as sensors and an ESP8266 module (to provide a communication link) in conjunction with constant monitoring of physical characteristics allows for the tracking of physical characteristics over time [2].

The subject of this project is related to thrips, which cause a reduction in crop yield of 30% to 50% and how an appropriate application of the research will lead to a management strategy to deal with disease and insect infestations. A machine learning (ML) based analysis was applied in conjunction with Internet of Things (IoT) sensors to determine the direct connection between pest lifecycle events and important parameters of the farm environment (mainly temperature, relative humidity, ambient air temperature and soil moisture). When visually diagnosing onion problems with ML-based image processing technology, the diagnosis made was correct in 96.3% of cases. It is important that the change in the color and moisture content of onion leaves be tracked, in order to accurately determine when it is becoming diseased. Researchers have demonstrated that color changes, as measured with various imaging technologies (e.g., RGB, or Near Infrared [NIR]), can indicate the presence of bacterial and fungal infections. The research demonstrates the importance of incorporating color detection systems to aid in identifying disease, to allow for the rapid detection of color changes and surface texture before infection spreads.

A variety of factors such as temperature, light intensity, and humidity contribute to this. To address this, smart farming monitoring system based on a three-layer IoT architecture is developed for real-time environment control. The results showed that maintaining optimal greenhouse conditions reduces bolting from 16.7% in open fields to 3% [6, 7]. Deep learning (DL) is being used more and more in recent research for automated pest and crop disease diagnosis. With explainable AI tools like Grad-CAM++ and LIME supporting mobile and web-based deployment, lightweight CNN models in conjunction with image augmentation approaches like Real-ESRGAN and MobileNetV2 have reached over 97% accuracy [18, 19]. With accuracies above 90% across a variety of datasets, transfer learning-based CNNs, Mask R-CNN, and YOLO variants have proven to perform well in illness identification, severity assessment, and pest counts [20]. Customized CNNs and hybrid transformer–CNN architectures have also been shown to achieve high classification accuracy for potato and onion leaf diseases, reaching up to 99.14% and 98.75%, respectively [21–23].

An IoT-enabled system is proposed for agricultural security system using an infrared camera, ultrasonic sensors and LoRa (Long Range) communications together. As animals moved about the field, an infrared camera would detect that motion and take a photograph of the animal; the photo would then be sent to a server for analysis via YOLOv8 object recognition software. Once the processing of the photo was complete, the information from YOLOv8 was sent to Firebase Cloud Messaging (FCM). Optimized YOLOv8 model attained a precision of 97%, recall of 96% and accuracy of 96% [24].

Many recent works have incorporated the usage of IoT technologies for the purposes of monitoring onions. However overall, these works have focused on either environmental sensing or using imagery as the sole method for detecting diseases. The proposed framework will differ from existing methods by also combining: Environmental Monitoring, CNN for disease classification on imagery; confidence weightings using multi-model fusion to increase reliability; using Grad-CAM for model explainability to increase interpretability of predictions; and using IoT technology in real-time at low cost for both spoilage detection across different storage conditions and for allowing remote access to view and monitor your crop. Table 1 presents the comparison of onion monitoring with the proposed framework.

3. Methodology

The research proposes a system to detect spoilage in onions in real-time by using IoT sensory data, DL classification of images, multimodal decision fusion and XAI. The structure of the proposed system includes three main components: (i) IoT sensing and collection of environmental data, (ii) multimodal processing and classification of onion image data using simple CNN, and (iii) evaluation, explainability, and generation of alerts.

Table 1. Comparison of recent onion monitoring studies with the proposed framework

Reference	IoT Sensors	CNN/DL	Multimodal Fusion	Explainable AI	Real-time IoT	Main Limitation
[1]	✓	✓	✗	✗	✓	No multimodal fusion or explainability
[2]	✓	✓	✗	✗	✓	Sensor and image data are not confidence fused
[4]	✓	✓	✗	✗	✗	Focused on disease detection without integrated deployment
[6, 7]	✓	✗	✗	✗	✓	Environmental monitoring only
[18–20]	✗	✓	✗	✓	✗	Image-based diagnosis only
[21–23]	✗	✓	✗	Partial	✗	No IoT sensing or multimodal decision fusion
Proposed	✓	✓	✓	✓	✓	Unified low-cost multimodal IoT framework

The use of sensors to continuously monitor product storage conditions and to capture images of onion products is used as the basis for computing and combining the classification results from environmental sensing data with the classification results from CNN classifiers according to a confidence weighted fusion technique in order to determine the product's final spoilage condition.

The work establishes three layers, the overall architecture of the system is made up of: a sensing layer, processing layer, and application layer. On the sensing layer, electrical signals generated by environmental parameters such as temperature, humidity, gas concentrations, and moisture are transmitted by an ESP32 microcontroller, which is connected to IoT sensors in the field. The processing layer then pre-processes these signals and sends them to a cloud platform, where a CNN performs classification based upon these input data. A multimodal fusion module combines the outputs from CNN prediction with the conditions scores generated from sensor data. Lastly, an application layer has mobile application interfaces that allow for visualization of these classifications and environmental conditions for the purpose of remote viewing and notification of readiness. The overall workflow of the three-layer architecture is illustrated in Figure 1.

3.1 Module 1: IoT Sensing and Data Acquisition

Considering a multi-sensor data acquisition, where T is the temperature sensed by DHT11 temperature sensor, H is the relative humidity sensed by surface and moisture sensor, G_3 and G_4 corresponds to MQ3 and MQ4 gas sensor readings, C represents the color features in the onion sensed by TCS3200 sensor, and M denotes the moisture level. The multi-sensor observation vector with respect to time t is represented as in Eqn. 1. Detailed sensor specification is listed in table 2.

$$S(t) = \{T(t), H(t), G_3(t), G_4(t), C(t), M(t)\} \quad (1)$$

Outputs from the DHT11, MQ3, MQ4, capacitive moisture sensor and TCS3200 color sensor were captured at a capture frequency of 1 Hz. Prior to feature extraction, all sensor channels had a moving average filter with a window length of 5 samples applied to attenuate high frequency noise and transient fluctuations in the output.The parameters are selected based on the strong correlation with onion spoilage mechanisms. Temperature and humidity influence microbial growth and enzymatic activity in onions, gas emissions serve as biochemical indicators of fermentation and anaerobic degradation, the visual symptoms are represented by color variations such as discoloration, mold formation and fungal depositions. By combining these parameters, the system captures early-stage biochemical changes and visible deterioration symptoms.

Before the experiments are conducted, all sensors were made stable and calibrated as per their Testing Storage Conditions (TSC). Because this framework uses a classification method based on process data instead of direct measurements of gas concentration, threshold values were developed from the experimental data after they were calibrated. The threshold values were then used when fusing the decisions made with multiple types of general sensor measurements.In the initial preprocessing stage, the noise and transient variant in the sensor signal is eliminated using a moving average filter. The equation governing is expressed in Eqn. 2.

$$\widehat{S_i}(t) = \frac{1}{N}\sum_{k=0}^{N-1} S_i(t-k) \quad (2)$$

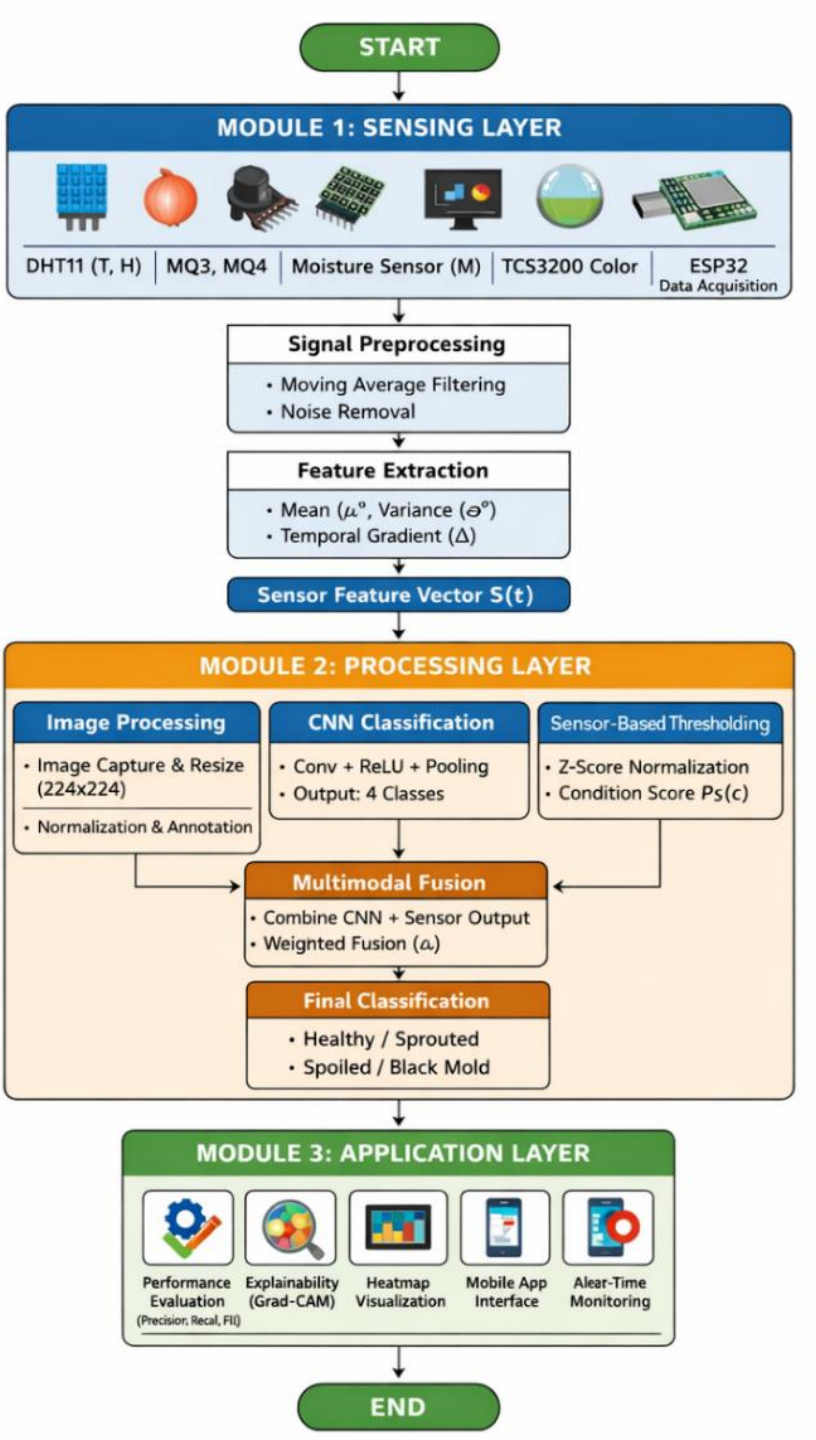


Figure 1. Flowchart of the proposed architecture.

In the above equation, N represents the window length and S_i is the i^{th} sensor channel.

From the filtered output signal, statistical features are extracted over a sliding window method.

$$\emptyset_i = \{\mu_i, \sigma_i^2, \Delta_i\} \quad (3)$$

Where μ_i is the mean, σ_i^2 is the variance and Δ_i denotes the temporal gradient capturing the rate of change, which is essential for determining the spoilage progression in the onions. Data collection for all sensors occurred prior to experiments with calibration occurring under: procedures were followed for calibrating the 1800 samples collected from each sensor in a steady state condition of storage of healthy onions. The mean (μ) and standard deviation (σ) of baseline data were determined from calibration data and used to statistically normalize the collected data and estimate an adaptive threshold. Prior to commencing an experimental trial recalibration were undertaken. Statistical features were extracted by using a sliding window of 30 samples (50 % overlap) for each sensor channel.

Table 2. Sensors specification

Sensors	Specification
DHT11	Range of Temperature: 0 to 50 °C (accuracy of ±2 °C) Range of Humidity: 20–90% RH (accuracy ±5%) Voltage range for operation: 3.3V to 5V One reading per second, or 1Hz, is the sampling rate.
ESP32	CPU: Tensilica LX6 dual-core with a maximum speed of 240 MHz 520 KB of SRAM Bluetooth 4.2 and Wi-Fi (802.11 b/g/n) connectivity 34 programmable pins make up GPIO. 3.3V is the operating voltage.
Surface and Moisture	3.3V to 5V is the operating voltage. Analog or digital output Area of Detection: Coated Probes
MQ3	Analog gas sensor for ethanol/VOC detection Operating voltage: 5 V Analog voltage output Heater power: approximately 800 mW.
MQ4	Analog methane/VOC sensor used for relative gas concentration measurement The operating voltage is 5V. The output is analog voltage. LPG, alcohol, smoke, and CH4 sensitivity
Color	Detection Range: Full RGB Output: Frequency, a measure of color intensity Light Source: Onboard white LEDs The operational voltage ranges from 3V to 5V.

Within every window, the mean, variance and temporal gradient for each sensor channel was calculated and provided a means of describing the environmental conditions and progression of spoilage of onions. The gradient feature is crucial for detecting early spoilage onset, as rapid increase in gas emissions, moisture, or temperature often lead visible deterioration. With the feature level representation of the sensor data, the system can easily differentiate between a "good" looking onion and a "bad" looking onion than simply using the raw sensor data from the sensor alone. In the system's initialization phase, all of the sensors were calibrated to eliminate any baseline drift from the sensor. Therefore, it is required that the sensor is recalibrated on a periodic basis in order for the sensors to maintain consistent levels of performance over long periods of time. The ESP32 microcontroller, pre-processing and aggregating the sensor data, performs this aggregation. Also included in the aggregation process are filters, feature extraction, and feature fusion.

An ESP32 microcontroller continuously acquired sensor data throughout the storage of onions. Temperature and relative humidity were acquired with a DHT11 sensor at a sampling frequency of 1 Hz, gas sensors, MQ-3 and MQ-4, acquired the concentration of ethanol and methane that would be released during the microbial decomposition of the onions. The moisture of the surface of the onions was obtained with a capacitive moisture sensor at the surface of the onions and the color intensity at the surface of the onions was obtained with a TCS3200 color sensor. The sensors' readings were merged together then sent to the processing unit at specific time intervals for analysis.

The extracted statistical features are normalized using z-score normalization. Eqn. 4 gives the mathematical expression of z-score normalization. x_i represents observed feature value and μ_i and $\sigma_i$are the mean and standard deviation obtained during calibration respectively. This step ensures that all the features are scaled before threshold-based confidence estimation.

$$Z_i = \frac{x_i - \mu_i}{\sigma_i} \quad (4)$$

3.2 Module 2: Multi-Modal Data Processing and Classification

3.2.1 Image Processing

Supermarkets and vegetable markets have provided the data directly. The dataset of onion photos was taken using a smartphone camera in real-world settings with different lighting, backgrounds, and angles.

For the purpose of training the model, images with resolutions ranging from 8 MP to 48 MP were obtained and then reduced to a consistent size (224 × 224 pixels). Different onion samples with differences in color, texture, size, and flaws are included in the dataset.

A total of 2,098 RGB "onions" were taken in real-life storage conditions with smart phones to create a dataset. Images of onions were captured from different views and with different lighting to create a realistic storage scenario. Each image was classified into one of four categories (healthy, sprouted, spoiled or black mold) by hand based on observable characteristics of spoilage. Each image represents a unique onion specimen and only one image is acquired for each onion bulb. Hence, explicit specimen identifiers are not required for partitioning, as each image represents a single specimen. To provide greater variation and model how the onions will be stored in reality, the dataset of images was taken from different angles, orientations and lighting conditions of multiple samples of onions. This partitioning method creates additional variability in the data set. The entire dataset is split into two subsets of images for CNN model training (80% of images) and testing (20% of images). Table 3 presents the total number of specimens acquired along with the training and testing split.

Table 4 presents the empirical sensor response range obtained from the experimental dataset and they are interpreted as reference values rather than deterministic classification threshold. The experimental results reveal that sensor responses corresponding to later onion spoilage phases are partly congruent with one another, a result of the non-stationary nature of gas emissions and environmental parameters. Hence, the separate analysis of sensor data does not allow to distinguish all onion conditions with certainty. To address the limitation, the proposed method utilizes a confidence-weighted multimodal fusion method whereby condition scores from sensors are integrated with class probabilities derived from CNN model. The ranges in table 4 are utilized for initial sensor assessment purpose and are not implemented as fixed decision points in classifying onion states. The moisture sensor and TCS3200 color sensor are used as complementary sensing modalities during multimodal fusion. The results from the sensor are incorporated after feature extraction rather than through fixed empirical threshold ranges as the color discrimination is primarily performed by using CNN. The proposed IoT-based onion spoilage detection system, represented in figure 2, showcases the hardware model that was created based on the provided description.

Table 3. Data Description

	Total No. of Specimens	Training (80%)	Testing (20%)
Healthy	512	410	102
Sprouted	505	404	101
Spoiled	510	408	102
Black Mold	571	457	114
Total	2098	1679	419

Table 4. Empirical sensor response ranges observed for different onion conditions and employed as reference values for sensor confidence estimation

Onion Condition	MQ3	MQ4	Temperature (°C)	RH (%)
Healthy	18–20	45–50	18–21	58–64
Sprouted	30–59	121–169	22–24	65–71
Spoiled	107–170	278–359	25–29	72–81
Black Mold	171–222	256–277	30–38	82–88

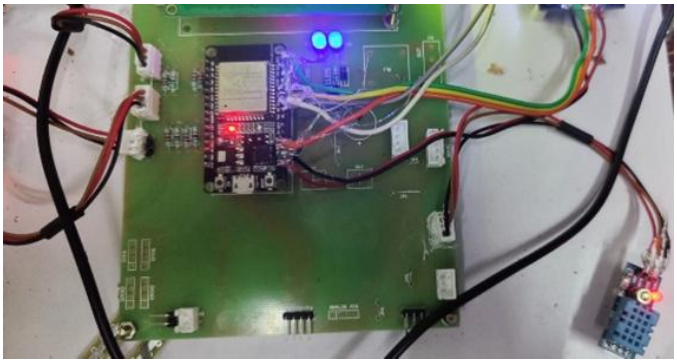


Figure 2. The developed prototype model.

The main processing unit or communication unit is identified by the green light, which has an active status; this is accomplished through an ESP32 microcontroller. The Printed Circuit Board (PCB) board integrates multiple connectors that link environmental sensors like humidity, gas, moisture and color sensor, and peripheral interfaces, through dedicated signal and power lines. The organized routing of wires and connectors reflect the modular design of the system, facilitating easier sensor replacement and scalability. The prototype includes on-board power regulation and status indication components, ensuring stable operation.

3.2.2 CNN Based Image Classification

The CNN architecture is made up of an input layer, two convolutional blocks, two max-pooling layer, one fully connected layer, and a Softmax output layer. ReLU activation will be used as an activation function after all convolutional layers. The architecture was designed so that it can deployed in an IoT-based storage-monitoring application. Table 5 shows the layers involved. The model comprises of 25.71 million trainable parameters, with majority contributed to the fully connected following the flatten layer.

Table 5. CNN Layers

Layer	Output Size	Trainable Parameters
Input	224 × 224 × 3	0
Conv1	224 × 224 × 32	896
MaxPool	112 × 112 × 32	0
Conv2	112 × 112 × 64	18,496
MaxPool	56 × 56 × 64	0
Flatten	200704	0
Fully Connected	128	25,690,240
Output	4	516

Let $I \in \mathbb{R}^{HxWx3}$ represent the RGB onion image. Non-linear mapping is learned through training CNN. The onion image is categorized into four classes. Each training class is calculated by the probability function. The equation governing is in Eqn. 5.

$$f_\theta : I \rightarrow p = [p_1, p_2, p_3, p_4] \quad (5)$$

Where p is the posterior probability of the four onion categories, Healthy, Sprouted, Spoiled and Black Mold, and θ is the learned network parameters. A CNN with 2 convolutional layers (ReLU activation), 2 max pooling layer, 1 fully connected layer (containing 128 neurons), and 1 Softmax output layer, has been built to classify images into four classes. Each convolutional layer output is mathematically expressed as in Eqn.6.

$$X^{(l)} = \sigma(W^{(l)} * X^{(l-1)} + b^l) \quad (6)$$

Where $*$ denotes the convolution operation, $\sigma(.)$ is the ReLU activation function, and $W^{(l)}$, b^l are the weights and bias of the l^{th}layer. The categorical cross-entropy loss is expressed in Eqn. 7.

$$\mathcal{L} = -\sum_{c=1}^{4} y_c log(p_c) \quad (7)$$

Where y_c is the ground truth label and p_c is the predicted probability for class c. To improve the generalization of the suggested work, data augmentation and k-fold cross-validation techniques are employed. The hyperparameter values of the CNN model is given as in table 6.

Table 6. Hyperparameter values of CNN

Parameter	Value
Optimizer	Adam
Learning Rate	0.0001
Batch Size	64
Epochs	30
Loss Function	Categorical Cross-Entropy
Dropout	Not used
Input Size	224 × 224 × 3
Random seed	42
Validation split	10%
Stopping rule	30 epochs
Model selection criteria	Lowest Validation Loss

3.2.3 Sensor-Based Adaptive Thresholding

The confidence values obtained from all the sensing modalities are normalized to generate sensor probability pertaining to four class represented by P_s. The probability vector represents the likelihood of each onion condition based on the sensor measurements and they are combined with the CNN probability vector using the confidence-weighted fusion method. Eqn.8 gives the probability vector calculation.

$$P_s = [P_H, P_{spro}, P_{spo}, P_{BM}] \quad (8)$$

Adaptive threshold is computed rather than fixed thresholds using z-score normalization, the mathematical expression for the computation of threshold is expressed as in Eqn. 9.

$$Z_i(t) = \frac{\hat{S}_i(t) - \mu_i}{\sigma_i} \quad (9)$$

Where μ_i and σ_i are baseline mean and standard deviation values obtained during calibration. A sensor-based condition score $P_s(c)$ for class c probability is estimated using Eqn. 10.

$$P_s(c) = \frac{1}{K}\sum_{i=1}^{K} \mathrm{I}(Z_i(t) \in \Omega_c) \quad (10)$$

Where Ω_c indicates the normalized threshold region for class c, K is the number of sensors, and $\mathrm{I}(.)$ is the indicator function, responsible for classification.To generate the sensor condition score, the normalized

readings generated from the temperature, relative humidity, MQ3, MQ4, moisture, and color sensors have been averaged together; moisture readings were considered in order to identify excessive surface wetness that may cause spoilage; the TCS3200 color sensor provided the amount of surface discoloration; the amount of discoloration indicates sprouting, fungal growth or black mold. The combination of the normalized values from each of the sensors is used to create an overall sensor condition score which is then merged with the CNN prediction during multimodal fusion.

3.2.4 Multi-Modal Decision Fusion

The final classification decision (Eqn.11) is obtained via confidence-weighted fusion of image and sensor outputs.

$$P_f(c) = \alpha . p_c + (1 - \alpha) . P_s(c) \quad (11)$$

Where $0 \leq \alpha \geq 1$ controls the relative influence of visual and sensor-based evidence. The predicted onion condition is given as in Eqn. 12.

$$\hat{c} = \arg\max_c P_f(c) \quad (12)$$

Using grid-search experiments the fusion parameter α is optimized with values from 0.1 to 0.9, and optimal value of α is found to be 0.7 based on the greatest macro F1-score achieved on the validation data set.

3.3 Module 3: Evaluation, Explainability, and alert generation

3.3.1 Performance Metrics

The model performance is evaluated using Eqn. 13-15.

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

$$F1\ score = \frac{2*Precision*Recall}{Precision+Recall} \quad (15)$$

An ablation study is conducted to quantify the contribution of the suggested model performance.

3.3.2 Explainable AI using Grad-CAM

Grad-CAM model is introduced to increase the interpretability. For target class c, the weight importance k^{th} feature map is calculated as in Eqn. 16. Let y^c represent the score for class c containing class labels as Healthy, Spoiled, Sprouted and Black-Mold.

$$\alpha_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (16)$$

The activation map is represented as $A^k$of the k^{th} feature channel in the final convolutional layer of the CNN. Z is the number of spatial locations in the given image and (i, j) is the pixel coordinates. Grad-CAM computes the importance weight α_k^c for each feature map by applying global average pooling over the gradients of the class score with respect to the activation map.

The Grad-CAM heatmap is mathematically expressed as in Eqn. 17.

$$L_{Grad-CAM}^c = ReLU(\sum_k \alpha_k^c A^k) \quad (17)$$

The ReLU activation function ensures that only the features that contribute to the predicted class are highlighted, eliminating irrelevant and negatively correlated regions. The resultant output image from Grad-CAM is up sampled to the original image resolution and overlaid on the input onion image to indicate the regions pertaining to the model's decision.The Grad-CAM images used in this study are only qualitative visualizations that describe how the model works and not for quantitative evaluation. Thus, the heatmaps used in this study do not make any claims to being quantitatively informative or accurate regarding whether the model's decisions were correct or significant. Sample Grad-CAM images are provided for correctly classified and misclassified onion samples.

3.3.3 Alert Generation and Deployment Constraints

The mobile app developed as part of the framework accesses and displays sensor readings using the Blynk IoT cloud service to allow for real-time decisions to be made and for remote monitoring and authorization to be performed. The app provides an easy-to-use dashboard to visually show the status of each sensor reading, any classification results made by the devices, and the status of the associated alerts. The ESP32 continuously collects environmental conditions such as temperature, relative humidity, MQ3 and MQ4 gas concentration levels, surface moisture, and color sensor data, and sends this data to the Blynk cloud service via Wi-Fi for remote monitoring.

The predicted condition of each onion, i.e., healthy, spoiled, sprouted, or moldy, is based on a coefficient-weighted combination of the condition scores of the onion determined by the CNN image classifier and the condition scores determined by each of the environmental sensors. This means that users can monitor their onions remotely without having to visit the location of the onions. A user interactive control interface is provided to allow users to turn their alert notifications on or off as they see fit. Alerts are sent whenever there is a high likelihood of spoilage in an onion according to the onion's condition by virtue of exceeding the pre-determined classification thresholds. Alerts would be sent to the mobile app from the Blynk platform via the cloud to allow for prompt identification and removal from storage of any onions that may have problems.

When an ESP32-CAM captures an image, the image is reduced to a resolution of 224 × 224 to keep the model input consistent when running the CNN. The system was validated in a controlled laboratory environment with an ESP32 sensor unit and using Blynk's IoT platform that provides real-time monitoring for alerts. The system performed successfully from end to end; however, the ability to evaluate performance fully under various network conditions (latency), differences in lighting, drift over time of sensors, partial occlusions, and mass storage environments from a commercial standpoint is outside the scope of this research and will be an important area for future research. The model parameter specifications for implementation can be found in Table 7. CNN inference is performed on host computer connected to ESP32-based IOT system, where ESP32 acquires images, collects sensor data and communicates to the cloud. The total trainable parameters is 25.71 million, 98.1 MB model size, 18 ms inference time per image and the average memory usage is 112 MB

4. Results

Data latency is evaluated to assess the performance of the suggested method in real-time. Latency is defined as the total time elapsed from image capture to result visualization in the application. Table 8 shows the results of data latency analysis.

According to experimental findings, end-to-end latency for spoilage detection and visualization is approximately 1.3 seconds. This is suitable for applications requiring near-real-time agricultural monitoring. It is estimated that 56% of the entire delay is caused by the uploading of pictures and the cloud inference process. From the user's perspective, the system provides timely feedback despite this without any noticeable delay. By coupling cloud-based deep learning inference with mobile interfaces and IoT devices, we can implement practical deep learning algorithms in post-harvest storage environments. Even more reductions in latency can be achieved by optimizing image resolution, using edge-based preprocessing, and using hybrid edge–cloud inference techniques.

Figure 3 shows a prototype of the proposed hardware prototype displaying sensor data on an LCD screen. In order to detect onion rotting, it is essential to monitor the temperature, humidity, and the gas concentrations regularly. Data is supplied to the ESP32 microcontroller by DHT11, MQ3, MQ4, and color sensor sensors. In order to provide customers with instant visual feedback, the software analyses data and displays the results. CNN model is trained using Adam Optimizer with a learning rate of 0.0001 with a batch size 64. The amount was trained for 30 epochs. To help avoid overfitting, early stopping was used when training. Multi-Classification categorical cross-entropy was applied as the loss function for this type of classification.

Table 7. Hardware and deployment specifications

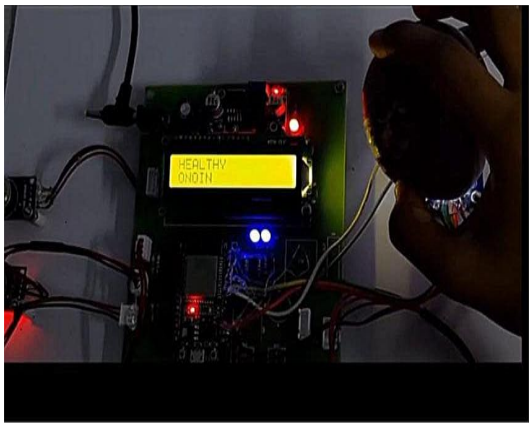
Parameter	Specification
Edge device	ESP32 Development Board
Cloud platform	Blynk IoT Cloud
Communication protocol	Wi-Fi (IEEE 802.11 b/g/n)
Sensor modules	DHT11, MQ3, MQ4, Capacitive Moisture Sensor, TCS3200
Image acquisition	ESP32-CAM
CNN input image size	224 × 224 pixels
Display	16×2 LCD (I2C)
Operating mode	Real-time monitoring

Table 8. Data Latency Analysis

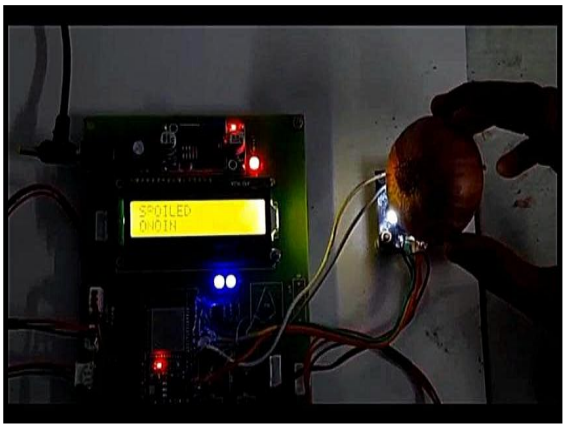
Process Stage	Average Latency (ms)
Image capture and preprocessing	180
Image upload to cloud server	420
CNN-based inference on cloud	310
Grad-CAM visualization generation	160
Result transmission to mobile app	230
Total end-to-end latency	1300 ms (≈1.3 s)



Figure 3. Proposed Prototype displaying the temperature, humidity and sensor values.



a) Result displaying healthy onion



b) Result displaying spoiled onion

Figure 4. The prototype displaying the result.

The prototype in operation during the onion quality classification process is shown in Figure 4(a and b). On the LCD, data is analyzed as it pertains to a healthy onion in Figure 4a, indicating perfect environmental conditions and no indications of rotting. An onion that has spoilt is evaluated in Figure 4b, and the LCD shows the diagnosis, demonstrating how the trained CNN model detects the spoilage condition with sensory inputs while working in the background. The purpose of this stage is to evaluate the model's ability to differentiate between healthy and rotting vegetables using real onion samples.

Using the sensor values shown in tables 9-12, it is possible to differentiate between onion health situations by combining temperature, relative humidity, and gas concentration levels (MQ3 and MQ4). Physiological and biochemical alterations induced by the progression of spoilage can be detected by the proposed sensing framework based on the observed patterns between sprouting, black mold, spoilt, and healthy onions. As shown in Table 9, MQ4 levels generally range between 45 and 50 parts per million, while MQ3 levels are ranges between 18 to 20 parts per million. Fresh onions that are metabolically active emit very little ethanol and methane due to the low gas concentrations. Its relative humidity levels (58-64%), along with its temperature values (18-21°C), are within ideal storage ranges, inhibiting the growth of microorganisms and inducing deterioration due to moisture.

In Table 10, spoilt onion trial values are shown. Readings of MQ3 and MQ4 are significantly increased in both samples. In contrast to MQ4 readings which rise to 278-359 Sensor Response (ADC Units), MQ3 readings rise to 107-119 Sensor Response (ADC Units), a significant increase. In addition to the increased gas emissions, onion deterioration is caused by anaerobic microbes and fermentation. The relative humidity of 72–81 percent and the temperature of 25–27°C provide ideal conditions for spoilage to accelerate.

Table 9. Healthy Onions

Trials	MQ3 Values (Sensor Response (ADC Units))	MQ4 Values (Sensor Response (ADC Units))	Temp (°C)	RH (%)	Result
1	18	49	19	64	Healthy
2	20	50	21	62	Healthy
3	19	46	21	62	Healthy
4	19	45	19	61	Healthy
5	19	50	20	61	Healthy
6	18	48	19	59	Healthy
7	19	47	20	62	Healthy
8	18	47	19	61	Healthy
9	18	49	18	58	Healthy
10	19	48	19	59	Healthy

Table 10. Spoiled onion

Trials	MQ3 Values (Sensor Response (ADC Units))	MQ4 Values (Sensor Response (ADC Units))	Temp (°C)	RH (%)	Result
1	107	290	26	76	Spoiled
2	129	283	26	73	Spoiled
3	176	359	27	80	Spoiled
4	143	278	25	77	Spoiled
5	154	295	27	78	Spoiled
6	111	298	27	79	Spoiled
7	138	292	26	76	Spoiled
8	166	287	26	72	Spoiled
9	142	291	27	78	Spoiled
10	179	353	27	81	Spoiled

Table 11. Black Mold Infested Onions

Trials	MQ3 Values (Sensor Response (ADC Units))	MQ4 Values (Sensor Response (ADC Units))	Temp (°C)	RH (%)	Result
1	173	278	37	84	Black Mold
2	189	266	30	88	Black Mold
3	177	256	33	86	Black Mold
4	193	287	37	83	Black Mold
5	172	257	38	85	Black Mold
6	171	269	34	82	Black Mold
7	201	291	31	87	Black Mold
8	197	288	38	88	Black Mold
9	222	284	37	86	Black Mold
10	199	293	36	88	Black Mold

Using gas sensors in real-time monitoring systems and detecting spoilage as early as possible is verified by the distinct difference between spoilt and healthy sensor profiles.

In all categories, the black mold-infested onions (table 11) have the highest sensor values. As MQ4 readings vary between 256 and 293 parts per million. The processed MQ3 sensor readings increased consistently with onion deterioration and were therefore adopted as discriminative features for spoilage classification.

These elevated gas concentrations reflect the powerful metabolic activity of the fungal community. The development of fungal colonies is closely related to high

relative humidity (82–88%) and high temperatures (30–38°C). A strong sensor response linked to black mold allows the project to distinguish advanced and dangerous deterioration phases from general spoilage, demonstrating its ability to detect advanced and dangerous deterioration phases.

Data from Table 12 indicate moderate sensor readings, both in healthy and spoilt circumstances.MQ3 values vary from 30 to 59 Sensor Response (ADC Units), indicating increased respiration and metabolic alterations caused by sprouting rather than decay, whereas MQ4 levels range from 121 to 169 Sensor Response (ADC Units). Both temperature range (22–28 °C) and relative humidity (67–75%) encourage sprouting activity without seriously degrading microorganisms. In order to enable condition-specific decision making and prevent misclassification as spoilt food, these data confirm that sprouting onions have a distinct sensor signature.

The IoT mobile application interface seen in Figure 5 was developed using the Blynk platform, enabling the user to view the categorization results from a distance. The interface's three main elements—Healthy Onion, Spoiled Onion, and Black Mold Onion—represent the many types of onion illnesses. By continuously updating with sensor data and classification percentages, the interface's dials give the user a clear, up-to-date picture of the state of the onions that are being stored. This app's remote monitoring and decision-making features would be especially helpful for farmers and supply chain managers who are not physically near the storage location.

A CNN-based onion classification model's confusion matrix is shown in the provided figure. The system divides onions into four classes: Black Mold, Sprouted, Spoiled, and Healthy. From the figure 6, the confusion matrix for the testing data is provided. Final classification is determined by weighting the CNN image prediction and condition score based on confidence in each type of input. Each type of evidence is used to determine the final classification is controlled by α, the weight controlling the fusion of visual and sensor-based evidence. Table 13 provides the validation performance across different α values. α is determined empirically by performing grid search experiments using a range of weights 0.1 to 0.9; the best value of α = 0.7 was identified as producing the highest classification performance using the validation dataset.

Figure 6 presents the confusion matrix of the testing set. A total of 419 images is considered for testing. A strong identification capability is demonstrated by the proposed model for different onion conditions based on the classification results. The majority of healthy onion samples are correctly identified, whereas a small number of sprouted and spoiled samples are incorrectly classified. This performance can be attributed to transient environment differences that may attributes to early spoilage or fungal activity.

The labels were incorrect for various onions that had spoiled; while the algorithm was able to achieve reasonable accuracy, it does not excuse the fact that errors will always happen. Most examples of black mold were misidentified as having spoiled onions by the algorithm; there was not much distinction between the two. The two types of spot defects and gases will assist in clarifying the cause of these misclassifications. The model has also produced results with very few misclassifications between fungal and non-fungal decay cases, indicating that the proposed methodology is a viable option for future use. The results further illustrate how various types of sensor fusion based upon CNNs resolve actual problems related to each of these onion samples. The model achieved an overall accuracy of 92.36%.

Table 12. Sprouted Onions

Trials	MQ3 Values (Sensor Response (ADC Units))	MQ4 Values (Sensor Response (ADC Units))	Temp (°C)	RH (%)	Result
1	33	167	27	74	Sprouted
2	30	159	26	68	Sprouted
3	37	121	27	67	Sprouted
4	43	134	27	73	Sprouted
5	47	138	28	75	Sprouted
6	53	156	24	72	Sprouted
7	59	161	22	67	Sprouted
8	56	128	28	68	Sprouted
9	48	169	27	69	Sprouted
10	39	166	23	74	Sprouted

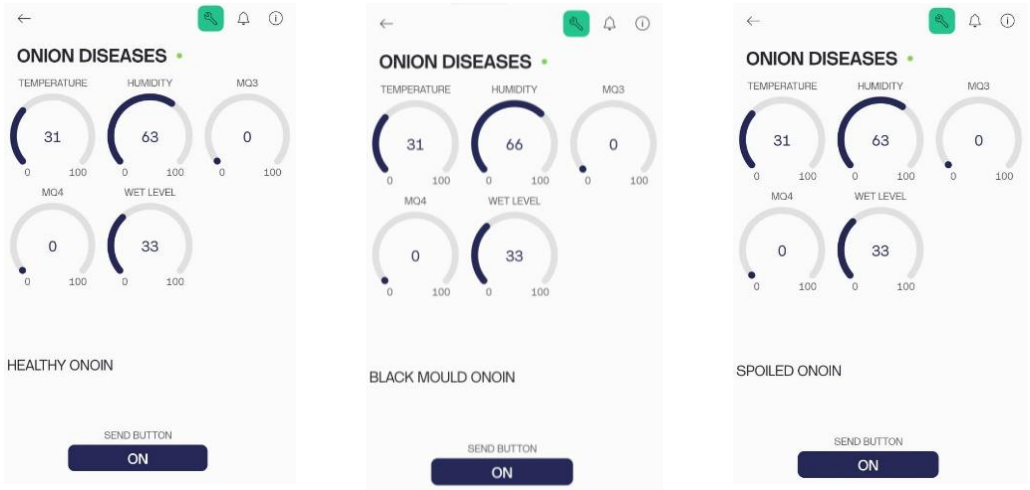


Figure 5. IoT App displaying the output.

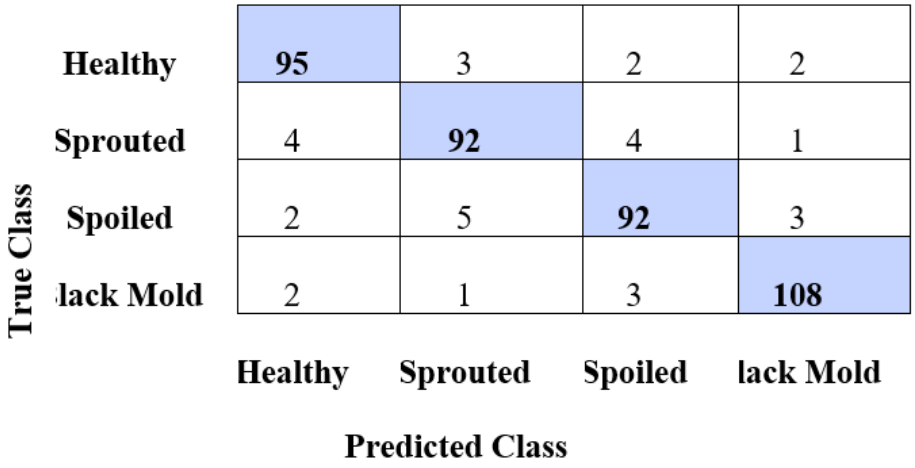


Figure 6. Confusion Matrix of CNN.

Table 13. Validation performance across different α values

α	Validation accuracy (%)	α	Validation accuracy (%)
0.1	88.7	0.2	89.6
0.3	90.5	0.4	91.3
0.5	92.0	0.6	92.8
0.7	93.5	0.8	92.9
0.9	91.8		

Table 14. Class-wise performance of the proposed CNN

Class	Precision	Recall	F1-score
Healthy	0.9223	0.9314	0.9268
Sprouted	0.9109	0.9109	0.9101
Spoiled	0.9107	0.9020	0.9060
Black Mold	0.9474	0.9471	0.9475
Macro Average	0.9229	0.9229	0.9229
Weighted Average	0.9232	0.9236	0.9233

Table 14 presents the class-wise performance of the proposed model. The model demonstrates balanced classification performance across all the classes, with black mold category achieving highest F1 score (0.9475), followed by healthy (0.9223). The sprouted and spoiled classes achieve reliable classification despite minor similarities. The macro and weighted average F1 sores of 0.9229 and 0.9223 confirms a robust and generalized model.

The CNN model is further evaluated based on 3-fold and 5-fold cross validation to ensure generalization. From table 15, 3-fold cross validation set, the model demonstrated consistent performance with minor variations in the precision, recall and F1 score. The validation loss across the set is low and stable. The model benefits from better learning, resulting in stable and better performance. In 5-fold validation, the input data is split into smaller subsets, thereby the model is evaluated rigorously. This type also showed consistent performance with fewer variations across all the evaluation metrics. This ensures that the model is not sensitive to data portioning and generalizes better for unseen data. Although the training occurs multiple times, the validation is more reliable in 3-fold validation. Overall, 3-fold offers faster computation, and 5-fold offers reliable and unbiased estimate of the model output. These results infer that the proposed CNN model achieves stable, and generalized performance making it reliable for real-world applications.

Results obtained from 3-fold and 5-fold cross validation indicate that both produce very similar average values, which should show that model performance is mostly stable regardless of how the data was split for use in training/testing. Results also show that the proposed CNN model has good generalization performance and is fairly insensitive to how the dataset was split.

Therefore, results from both types of cross validation primarily serve to demonstrate reproducibility of the proposed CNN model as opposed to providing an independent evaluation of the model.

Figure 7 shows the output of the Grad-CAM model. The purpose of this research was to provide qualitative assessments of how image region contributes to the decisions made by the CNN classification system by generating Grad-CAM heat map visualizations. The heat maps provide areas that are most influential in how the CNN made its decision, which helps to provide greater interpretive capabilities of its decision-making process. The Grad-CAM visualizations generated for this study are intended to provide qualitative explanations only and were not validated quantitatively against the expert-annotated regions of spoiling. The visualization makes it easier to see the onion's damaged areas. The creation of a heat map in this output makes identification simpler.

Table 15. K-fold cross validation of CNN model

K=3				
Fold	**Precision**	**Recall**	**F1-Score**	**Validation Loss**
Fold 1	0.918	0.932	0.924	0.190
Fold 2	0.912	0.928	0.920	0.205
Fold 3	0.921	0.935	0.927	0.185
Average	0.917	0.932	0.924	0.193
K=5				
Fold 1	0.918	0.932	0.924	0.190
Fold 2	0.915	0.930	0.922	0.198
Fold 3	0.920	0.934	0.926	0.187
Fold 4	0.913	0.929	0.921	0.202
Fold 5	0.919	0.933	0.925	0.189
Average	0.917	0.932	0.924	0.193

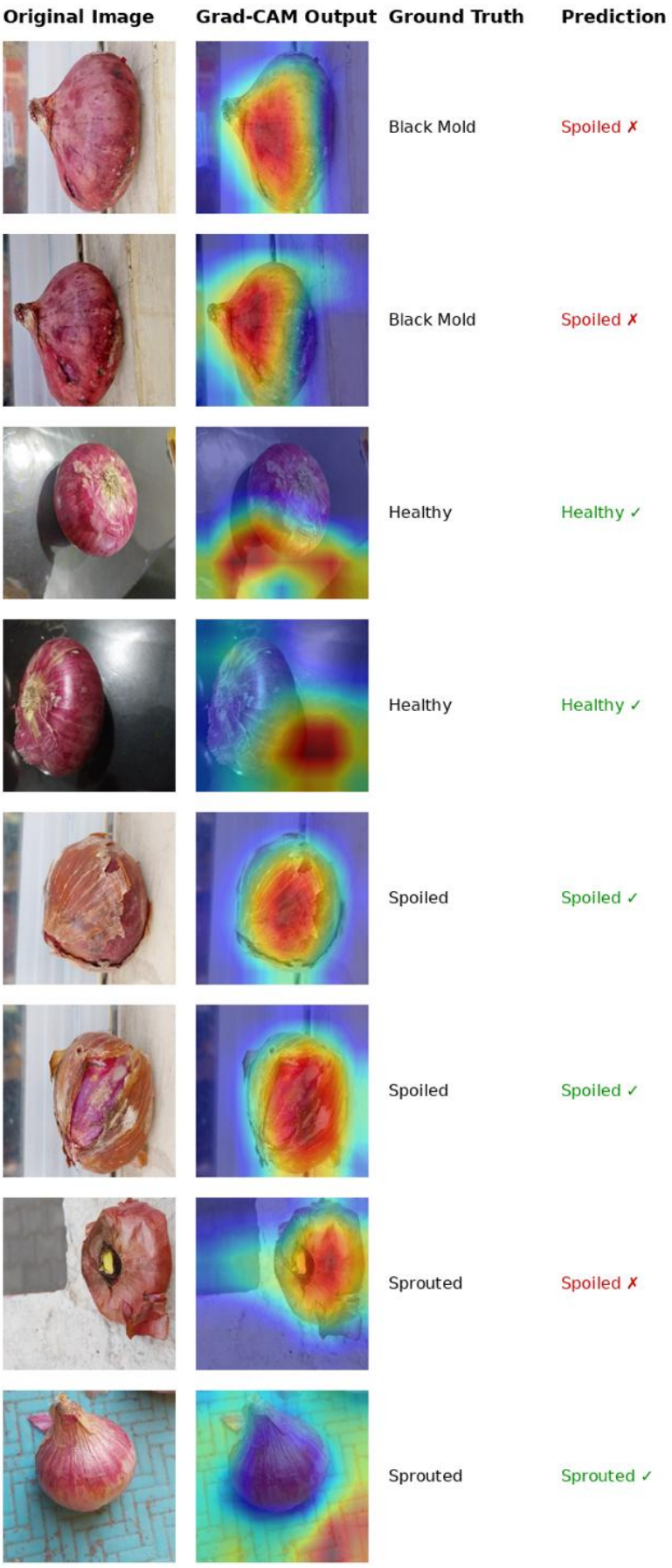


Figure 7. Grad-CAM output.

5. Analysis and Discussion

Agricultural countries (like India) face a large issue with regards to post-harvest loss for onions when they are stored (rotting) where the percentage (20%-40%) of rotted onion crops. Most of these losses are a result of microbial infection - specifically fungal diseases (Aspergillus niger). This fungus grows in high humidity and temperature environments and worsens with improper handling of the onions and not enough ventilation. A result of the traditional storage of onions is farmers do not have a way to continuously monitor biological/environmental conditions therefore resulting in substantial financial losses. The financial losses to farmers can also affect national food security and national market prices of onions and the impact the storage has on the supply of food for consumers/families.

The ESP32 microcontroller, which is a low-cost (dual core) microcontroller with Wi-Fi capability, will handle basic processing for the system's multi-modal sensor input to transfer that data to Blynk IoT for use on a global scale.Sensors include a capacitive moisture sensor that has been repurposed for surface wetness detection and the DHT11 for ambient temperature and humidity monitoring (range: 0–50 °C and 20–90% RH, with ±2 °C and ±5% accuracy, respectively). Two different types of gas sensors, MQ3 and MQ4, were used to identify volatile substances that signify the deterioration of onions. The sensor data collected from both sensors, along with environmental data and image features, were used to accomplish multimodal classification of onions. Even before obvious symptoms appear, these early-stage molecular markers provide prognostic insight into spoiling conditions. For remote monitoring, the data is simultaneously sent to the user's mobile device and shown locally on a 16x2 LCD with an I2C interface.

A CNN-based visual inspection module was created in addition to the sensor array, and it was trained using a dataset of annotated onion photos that were divided into groups for ruined and healthy onions. CNN has a set of two convolutional layers activated by the ReLU function each followed by two max pooling layers, one fully connected layer with 128 neurons, and one Softmax layer to classify 4 classes total. This architecture offers a balance between speed and accuracy when classifying data. A CNN architecture was chosen to tradeoff between high levels of accuracy in classifying objects, while having low computational complexity for use in an IoT application where real-time processing is necessary. The model was trained utilizing data augmentation approaches, such as rotation, brightness adjustment, and image flipping, to improve robustness under real-world storage fluctuation. With precision, recall, and F1-scores of 91.8%, 93.2%, and 92.4%, respectively, the CNN demonstrated good generalization and classification reliability.

Moreover, various multi-dimensional sensor signatures rather than a single dominant characteristic characterize onion circumstances, according to the combined analysis of Tables 4–7. While spoilt and black mold-affected onions display high gas concentrations associated with unfavorable environmental circumstances, healthy onions have modest gas emissions and steady temperature–humidity profiles. Sprouted onions have sensory characteristics that can be divided into three stages, which allows for intelligent classification models to classify the stage of deterioration in sprouted onions. The ability to validate the effectiveness of multi-sensor data fusion for monitoring onion quality is accomplished through stratification of the sensory data. The system is powered by an LM2596 buck converter. This allows for the stable operation of the system to use a wide range of power sources such as rechargeable batteries, solar panels, or a variety of grid-supplied power. As a result, this design allows for reliable deployment in remote and resource-scarce environments.

The suggested multimodal fusion framework has a great advantage during periods of transitional spoilage where environmental changes and volatile gas emissions occur before there are observable signs of spoilage on the onion’s outer skin. The image-based classification system (CNN) is effective at distinguishing classes visually, however, the addition of data from sensors provides a greater level of confidence in the final decision because it includes data from the environment in relation to spoilage. Therefore, this approach decreases the chances of producing an incorrect prediction due to the difficulty to accurately predict spoilage when the environment is volatile.

5.1 Ablation Study

An ablation study is conducted to evaluate the effectiveness of the proposed approach. In the first case the data from the sensor alone is used, the second case uses CNN model alone, and the third approach employs the fusion method using the data from the sensor and categorizing using CNN. Table 16 presents the results obtained using these input combinations. This ablation study shows that the CNN model is much better at classifying onion deterioration than an approach that uses only sensor data, because it measures visual features of the onions' state. The overall classification performance of the multimodal framework, when measurements from the environmental and gas sensors were combined with the CNN output, was enhanced by providing complementary information, especially when there are not enough visual symptoms to differentiate the onions reliably. Therefore, the multimodal framework provides more robustness than either of these two sensing modalities when used alone.

Table 16. Ablation study

Class	Sensor Only (F1)	CNN Only (F1)	Proposed Fusion (F1)
Healthy	0.81	0.93	0.95
Sprouted	0.75	0.91	0.94
Spoiled	0.77	0.91	0.93
Black Mold	0.80	0.95	0.97
Macro Average	0.78	0.92	0.95

Table 17. Comparison of fusion strategies

Fusion Method	Accuracy (%)
Equal-weight averaging	91.5
Majority voting	90.9
Maximum confidence	91.8
Confidence-weighted fusion (α = 0.7)	92.4

With the multimodal fusion strategy, a notable performance improvement is identified in distinguishing between different products during transitional spoilage conditions when visual symptoms on their own were not enough to consistently distinguish between them. By using environmental sensor measurements along with the CNN's predictions, earlier biochemical changes and gas emissions are tracked than were visually evident, which helped to increase the overall robustness of our results and reduce misclassification rates.

5.2 Fusion Comparison

The proposed fusion (confidence – weighted) model is compared with equal-weighting averaging, and majority voting. The proposed model achieved better accuracy compared to the other two models. Table 17 presents the experimented results.

6. Conclusion

Calibrated devices for example ESP32 were paired with sensors including but not limited to Color Sensors, MQ3/MQ4 Sensors and DHT11 sensors, providing accurate environmental monitoring, which is done Remotely and in Real-Time via the Blynk IoT platform. Quantitative validation of the CNN model produced a low, validation loss of 0.19, along with a precision rate of 91.8%, a recall rate of 93.2%, and an F1 score of 92.4%, indicating the accuracy of the CNN model. Furthermore, a confusion matrix demonstrated the validity of the model's classification of Onion Types. The Prototype, Mobile Apps, and Grad-CAM visual outputs provide evidence of the system's functionality, and user-friendly interface. Overall, this method is cost-effective, scalable, energy-efficient and will provide farmers with valuable information while identifying concerns earlier in the production process; thus, promoting Sustainable Agriculture through the reduction of post-harvest loss.

Future work will include performing a detailed analysis of the overall class performance of the fusion system, comparing the fusion system to other decision fusion methods, and developing an adaptive method for determining the fusion weight instead of using the empirical method through grid search.

7. Limitations

In this work, partitioning of data into groups is done at an image level. Given that the images were acquired under different acquisition conditions, which contribute to the diversity of the images, it would have been preferable for the partitioning to be done at the specimen or subject level to create an independent set of training and testing samples. Future studies will use this approach to reduce the chance of data leakage and to provide a more robust evaluation of the ability of the model to generalize to new images.

References

[1] D. Spurthi, P.C. Swathi, Yerra, S.C., Balaji, M.D. Pai, D., Early Detection of Onion Spoilage Utilizing IoT and AI during Storage and Transportation. International Journal of Innovative Science and Research Technology (IJISRT), 9(4), (2024).

[2] K.A. Shah, M.U. Sri, B.V. Goud, K. Mannem, (2023) Two-Fold Spoiled Onion Detection using Soft Computing and IoT. In 2023 International Conference on Inventive Computation Technologies (ICICT), IEEE, Lalitpur, Nepal.

https://doi.org/10.1109/ICICT57646.2023.10134248

[3] R. Ali Eltabey, Utilizing Digital Technologies to Ensure Food Safety. International Journal of Artificial Intelligence and Emerging Technology, 6(2), (2023) 42-65.

[4] S. Banerjee, K. Sarda, S. Khandelwal, A. Gajarushi, S. Gawande, R. Velmurugan, M.S. Baghini, (2024) IoT-based Sensing System for Thrips Pest and Disease Management in Onion Crop. In 2024 IEEE Applied Sensing Conference (APSCON), IEEE, Goa, India. https://doi.org/10.1109/APSCON60364.2024.10466190

[5] V.B. Singh, N. Sharma, S. Singh, A review of imaging techniques for plant disease detection. Artificial Intelligence in Agriculture, 4, (2020) 229-242. https://doi.org/10.1016/j.aiia.2020.10.002

[6] Z. Khan, M.Z. Khan, S. Ali, I.A. Abbasi, H.U. Rahman, U. Zeb, H. Khattak, J. Huang, Internet of Things-Based Smart Farming Monitoring System for Bolting Reduction in Onion Farms. Scientific Programming, 2021(1), (2021) 7101983. https://doi.org/10.1155/2021/7101983

[7] P. Gavali, A. Gend, D. Birari, R. Garad, IOT Based Smart Farming Monitoring System for Bolting Reduction in Onion Farms, 8(3), (2023) 86-90. https://ijsdr.org/papers/IJSDR2303018.pdf

[8] K.N. Vhatkar, A.B. Kathole, J. Katti, G.J. Dharmale, An Innovative Onion Disease Classification Framework Using Smart IoT Assisted Multi-Scale Attention-Based ResNeXt with SPP via Modified Cat–Mouse Optimizer. International Journal of Image and Graphics, 26(08), (2025) 2750028. https://doi.org/10.1142/S0219467827500288

[9] A. Popa, M. Hnatiuc, M. Paun, O. Geman, D.J. Hemanth, D. Dorcea, L.H. Son, S.I. Ghita, An Intelligent IoT-Based Food Quality Monitoring Approach Using Low-Cost Sensors. Symmetry, 11, (2019) 374.

[10] E.G.A.M. El-Dawy, M.A. Hussein, S. El-Nahas, Description and Management of Aspergillus Section Nigri Causing Post-Harvest Bulbs Rot of Onion. Scientific Reports, 14, (2024) 6076. https://doi.org/10.1038/s41598-024-53849-9.

[11] M. Labanska, S.V. Amsterdam, S. Jenkins, J.P. Clarkson, J.A. Covington, Preliminary Studies on Detection of Fusarium Basal Rot Infection in Onions and Shallots Using Electronic Nose. Sensors (Basel, Switzerland), 22(14), (2022) 5453. https://doi.org/10.3390/s22145453

[12] M. Imran, H. Kang, S. Lee, E. Kim, H. Park, S. Oh, Current Trends and Future Prospects in Onion Production, Supply, and Demand in South Korea: A Comprehensive Review. Sustainability, 17(3), (2025) 837. https://doi.org/10.3390/su17030837

[13] T. Nivethitha, P. Vijayalakshmi, J. Jaya, S. Shriram, (2022). A review on coconut tree and plant disease detection using various deep learning and convolutional neural network models. In 2022 International conference on smart and sustainable technologies in energy and power sectors (SSTEPS), IEEE, Mahendragarh, India. https://doi.org/10.1109/SSTEPS57475.2022.00042

[14] P. Bhavesh, N. Ramakrishnan, (2025) An Interpretable and Efficient Deep Learning Framework for Leaf Disease Detection in Agriculture. In 2025 IEEE 6th India Council International Subsections Conference (INDISCON), IEEE, Rourkela, India. https://doi.org/10.1109/INDISCON66021.2025.11253779

[15] C. Harishankar, N.S. Anoop, K.S. Niranjana, A. Biju, V. Venugopal, S. Vekkot, (2024) An Explainable Hybrid Learning Model for Indian Food Image Classification. In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, Kamand, India. https://doi.org/10.1109/ICCCNT61001.2024.10724730

[16] R.R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. International Journal of Computer Vision, 128, (2016) 336–359. https://doi.org/10.1007/s11263-019-01228-7

[17] S. Manoj, M. Sathyanarayana, S.E. Neelagund, AI-Assisted Smart Food Packaging: Current Technologies, Challenges, and Future Opportunities. International Journal of Science, Strategic Management and Technology, 02(05), (2026). https://doi.org/10.55041/ijsmt.v2i5.130

[18] R.C. Balasooriya, E.L.A.Y. Adithya, M.M.S. Gunarathne, T.C.D. Silva, S. Lokuliyana, P. Wijesiri, (2025) An Explainable Deep Learning Framework for Coconut Disease Detection Using MobileNetV2, Super-Resolution, and Grad-CAM++. In 2025 International Conference on Data Science and Its Applications (ICoDSA), IEEE, Jakarta, Indonesia. https://doi.org/10.1109/ICoDSA67155.2025.11157166

[19] G.M. Abhishek Drusti S. Shastri, Amit Kumar Yadav, S.V. Aisiri, Sugarcane Disease Detection Using Deep-Learning and LIME. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 8(4), (2025) 6845-6850.

[20] S. Vidhanaarachchi, J.L. Wijekoon, W.S.P. Abeysiriwardhana, M. Wijesundara, Early

Diagnosis and Severity Assessment of Weligama Coconut Leaf Wilt Disease and Coconut Caterpillar Infestation Using Deep Learning-Based Image Processing Techniques. IEEE Access, 13, (2025) 24463–24477. https://doi.org/10.1109/ACCESS.2025.3537664

[21] M.J. Hoque, M.S. Islam, An Explainable and Lightweight CNN Framework for Robust Potato Leaf Disease Classification Using Grad-CAM Visualization. Applied AI Letters, 7(1), (2026) e70011. https://doi.org/10.1002/ail2.70011

[22] B.B. Vikhe, P.R. Patil, P.R. Bhaladhare, A. Potgantwar, Image-Based Onion Leaf Disease Identification using a CNN-Transformer Hybrid Approach. International Journal of Image and Data Fusion, 16(1), (2025) 2561938. https://doi.org/10.1080/19479832.2025.2561938

[23] S.N. Zaki, M. Ahmed, M. Ahsan, S. Zai, M.R. Anjum, N.U. Din, Image-Based Onion Disease (Purple Blotch) Detection Using Deep Convolutional Neural Network. International Journal of Advanced Computer Science and Applications, 12(5), (2021) 448-458. https://doi.org/10.14569/ijacsa.2021.0120556

[24] P. Singh, R. Krishnamurthi, IoT-Based Real-Time Object Detection System for Crop Protection and Agriculture Field Security. Journal of Real-Time Image Processing, 21(4), (2024) 106. https://doi.org/10.1007/s11554-024-01488-8

Authors Contribution Statement

Both authors contributed equally to this work, and both have read and agreed to the published version of the manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License

© The Author(s) 2026. The text of this article is open access and licensed under a Creative Commons Attribution 4.0 International License.