Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Identification of Maize Seed Quality Using Hybrid Deep-Learning Techniques and Multimodal Sensing Technologies

Dr. Ganpat Joshi[1], Dr. Ganga Singh Chouhan[2]

[1]Faculty of Engineering and Technology (FOET), Madhav University Abu Road Sirohi, Rajasthan, India

[2]Faculty of Commerce and Management, Madhav University Abu Road Sirohi, Rajasthan, India

[1]Shiv.joshi322@gmail.com,[2]ganga.singh@madhavuniversity.edu.in

Abstract:

Reliable recognition of maize seed quality is essential for increased crop production and food safety as well as for automatic seed certification. Traditional visual inspection and single sensor systems can be subjective, laborious to perform, and incapable of identifying subtle or internal defects. The deep learning approach However, current methods have limitations as the segmentation of maize seed, especially bad quality maize seed, is affected by noise and low contrast as well as orientation, meaning that it will be difficult to extract features. The procedure builds on the combination of RGB imaging, hyperspectral and NIR sensing to record complementary spatial, spectral, and biochemical properties of maize seeds. Modality-specific feature extractors are used, spatial features learnt by convolutional neural networks (CNNs) and wavelength dependent information modelled by spectral deep models. We propose a hybrid CNN Transformer model that captures both local texture patterns and global contextual dependencies. In addition, attention-based multimodal fusion mechanism is proposed to automatically and dynamically fuse redundant features across various sensors. Comprehensive experiments on a designed multimodal maize seed dataset show that the proposed framework considerably outperforms traditional machine-learning-based methods, single-modality deep models, and baseline fusion strategies in terms of classification accuracy, F1-score and robustness against illumination and noise perturbations. Explainable AI techniques are used to visually display discriminative regions and crucial spectral bands, providing interpretability for agricultural users. Results indicate strong potential of the proposed system for real-time application at automated seed grading and sorting plant, and it

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

can be considered as an effective, scalable, and reliable tool for intelligent assessment of seed quality in precision farming.

Keywords: Maize seed quality, Hybrid deep learning, Multimodal sensing, Hyperspectral imaging, Attention-based fusion

1. Introduction

Maize is one the world's most important crops, used as human food, animal feed and industrial raw material. The quality of maize seeds is an important factor affecting germination rate, uniformity of emergence, yield stability and resistance or tolerance to biotic and abiotic stresses. Therefore, rapid determination of the quality degree of maize seed is indispensable in all links along the agriculture value chain: seed breeding, processing, certification, storage, and shipment. The increasing dependency on yield and climate resilient crops, led to the urgency of intelligent, unbiased, and high-throughput quality testing systems for seed.

1.1 The strategic importance of identifying seed quality in the maize supply chain

Seed quality indicates the agronomic performance of maize crops and the economic feasibility of farming systems is determined directly by seed. Seeds with undesirable qualities, such as mechanical injury, fungal and insect infestation or physiological immaturity, may result in low seed germination percentage and reduced yields. Industrially, standardized seed quality determination is essential for grading, automated sorting and traceable certification paving the way to increased farmer confidence and compliance with regulations. Precise quality evaluation is a necessity in scale seed companies to reduce post-harvest losses and provide optimum product quality.

1.2 Limitations of Conventional Inspection and Single-Sensor Approaches

The conventional seed quality assessment relies heavily on manual examination and deleterious laboratory testing. These methods are subjective, time consuming and not good for real-time or bulk processing. Further, human visual inspection cannot identify hidden internal defects or biochemical changes that are important for the determination of seed

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

viability. Single-sensor based inspection systems (and specifically RGB), are characterised by a poor discriminative power, as they do not manage to exploit the spectral, structural and compositional properties at once. Therefore, the systems have lower stability in terms of illuminations, seed species and environment changes.

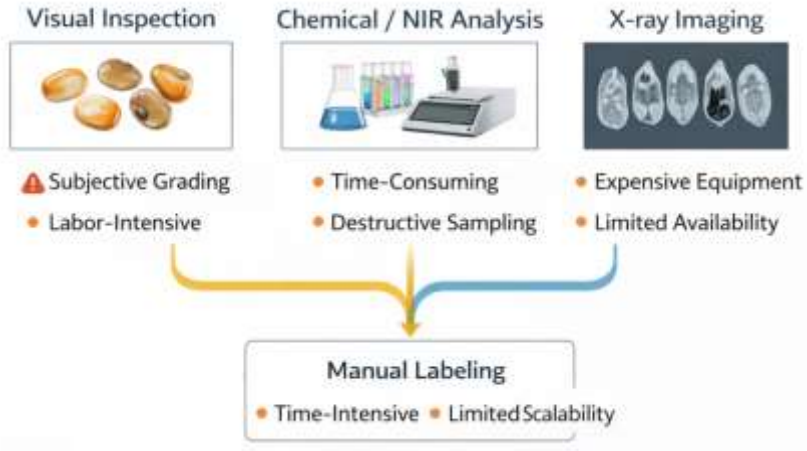

Figure 1. Illustration of limitations in conventional maize seed quality identification methods.

The figure reveals limitations of visual manual inspection and single-sensor imaging methods in terms of: human subjectivity, slow throughput rate, poor ability to detect sub-surface defects, lack of robustness with varying seed and lighting conditions. These restrictions spark the demand of intelligent, multimodal and automatic seed quality analysis systems.

1.3 AI-Powered Sensing in Automated Seed Grading

Artificial intelligence (AI) and sensor technology has recently promoted the reform of agriculture quality inspection. Deep learning based methods, particularly convolutional neural networks (CNNs), have achieved encouraging results for agricultural image analysis by automatically learning hierarchical feature representations from data. At the same time, non-invasive sensing approaches for hyperspectral or near infrared (NIR) imaging yield diverse spectral features that are correlated with chemical composition, moisture distribution and internal structural integrity. The intersection of AI and advanced sensing has given rise to data-driven, automatic and objective seed grading systems that perform better than the rule-based, traditional ones.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

1.4 Motivation for Hybrid Deep Learning and Multimodal Feature Fusion

Existing deep learning-based seed inspection approaches have achieved some promising results, however the majority of studies rely on unimodal data and homogeneous model architectures, which limits their capability to model the internal complexity of seed quality related attributes. These are all combined into deep learning-based hybrid architectures, in which convolutional networks are integrated with attention/transformer models, providing an efficient solution to simultaneously capture fine-grain local patterns as well as long-range context dependencies. In addition, by jointly aggregating the multi-modal features in RGB and spectral domains through concatenation feature fusion operation, the complementary cues can be effectively integrated to boost robustness, generalization and discriminative performance. These considerations inspire the hybrid deep learning architecture that integrates attention-based multimodal fusion in the context of maize seed quality recognition.

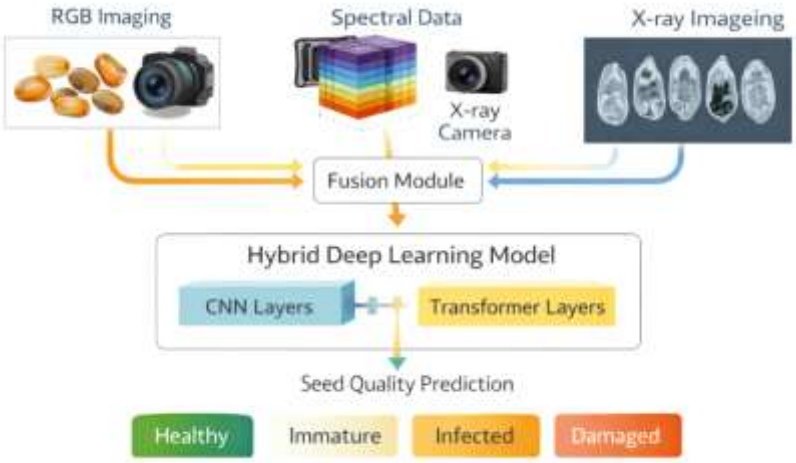

Figure 2. Conceptual framework of the proposed hybrid deep learning and multimodal sensing approach.

Harvesting external biological characteristics in RGB imaging and the internal structural properties (biochemical and hyperspectral/NIR spectral) are two completely different tasks. The modality-specific deep feature extractors are further connected with an attention-based fusion module, and thus robust identification of maize seed quality is achieved.

1.5 Research Objectives

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

The aim of this study is:

- To create a non-invasive, multimodal sensing system for holistic analysis of the quality of maize seed.
- A novel deep learning architecture that effectively integrates spectral information with spatial information is to be developed.
- To introduce an attentnionbasedcrossmodal fusion approach that adaptively fuses the heterogeneous sensor inputs.
- To perform extensive experimental analysis with respect to traditional machine learning models as well as contemporary deep learning baselines.
- To evaluate the general robustness of the developed framework for real-time and industrial seed grading systems.

1.6 Major Contributions

The main contributions of this work include:

- An overlaying system of RGB image-based solutions and spectral sensing technologies for maize seed quality identification."
- A new hybrid CNN-Transformer model that capture well small texture features and large relationships.
- An attention-based multimodal fusion model, which improves feature complementation and classification reliability.
- Comprehensive quantitative and qualitative comparisons, including ablation and robustness studies, show consistent improvement over state-of-the-art methods.
- Integrating methods from explainable AI to increase transparency and trust of model predictions for agricultural stakeholders.

Table 1. Comparison of traditional inspection methods, single-modality AI systems, and the proposed hybrid multimodal deep learning approach.

Approach Type	**Sensing Modality**	**Key Advantages**	**Major Limitations**
Manual inspection	Human visual	Low cost, simple	Subjective, slow, error-

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Approach Type	Sensing Modality	Key Advantages	Major Limitations
	assessment		prone, non-scalable
Laboratory testing	Destructive chemical/physical tests	Accurate, standardized	Time-consuming, destructive, costly
Single-sensor imaging	RGB camera	Non-destructive, fast	Limited feature representation, low robustness
Single-modality deep learning	RGB or spectral only	Automated, higher accuracy	Fails to capture complementary information
Proposed hybrid multimodal DL	RGB + spectral (HSI/NIR)	High accuracy, robustness, scalability	Higher system complexity (manageable)

2. Related Work and Literature Review

2.1 Old Methodology of Seed Quality Examination

Traditional corn seed quality control is based on visual examination, physical purity measurement and laboratory germination/vigor tests. These standards of practice are time consuming, analyst-dependent operations that cannot enter into high-throughput scanning modes. In addition, a number of major quality defects internal tissue damage in which insect infestation are not visible from outside and biochemical degradation caused during storage cannot be identified using only visual inspection. Such shortcomings have encouraged the development of non-destructive imaging and sensing technologies for revealing internal structural and compositional information present in seeds [3], [5].

2.2 Machine Learning of Seed Quality Discrimination

Previous attempts to automate seed grading have applied shape descriptors with color, texture and spectral indices as hand made features in combination with conventional classifiers such as support vector machines (SVM), random forest (RF) and partial least squares discriminant analysis (PLS-DA). The spectral-based approaches, such as NIR

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

spectroscopy and HSI-based methods, usually require a huge number of preprocessing treatments (e.g., MSC, SNV normalization and SG smoothing) and feature selection strategies including SPA (successive projections algorithm) and CARS (competitive adaptive reweighted sampling) before modeling training [2], [11]. Although these pipelines can work well under certain conditions, the generalization performance will be significantly degraded in the presence of variety shifts, illumination changes and sensor variations, which heavily rely on a large amount of domain-specific tuning [2], [16]. Recent research suggests that for NIR data, machine learning models can facilitate support seed germination and origin prediction, although stability under more general operating circumstances and over multiple datasets is a critical issue [12], [13]

2.3 Deep Learning techniques applied in agricultural inspection (CNN, ViT, LSTM)

The introduction of deep learning has significantly promoted seed inspection through the possibility to end-to-end learn representations directly from image and spectral data. Convolutional Neural Networks (CNNs) are extensively used for seed sorting and defect detection, such as deep learning models based on RGB found to be successful in large-scale germination prediction [3]. In relation to X-ray seed testing, CNN models have been trained for internal tissue quality (physiological) and integrity [4], and routine systems based on object detection networks, as is the case with YOLO and Faster R-CNN, have exhibited excellent prediction performance; good candidate solutions are also found in industry based processing pipelines [5].Additional deep embodiments Deep architectures can learn joint spectral-spatial representations for hyperspectral data cubes. Attention mechanisms based on transformers and hybrid models combining CNN and sequence processing (CNN-LSTM architecture) have been utilized to model long-range dependencies between wavelengths and global context information [1], [10], [15]. Together, these advances provide the motivation to investigate hybrid CNN-Transformer architectures for an improved maize seed quality evaluation.

2.4 Multimodal Sensing for Seed Quality Analysis

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

(a) **RGB:** It is very cost effective and fast method for external defect inspection, but it has limited sensitivity to internal or biochemical irregularities and strongly depends on illumination [3].

(b) **Hyperspectral/multispectral (VIS-NIR):** A hyperspectral imaging includes spatial and spectral information, which can help to detect fine structural damages and material differences. Examples of applications in maize seed analysis are the determination of coated kernel varieties and defect detection [1], [2]. In addition, HSI has demonstrated successful implementation in seed viability and vigor determination in other crops of the same family, indicating its capacity for evaluations related to physiological quality [11].

(c) **NIR spectroscopy**: NIR spectroscopy allows swift and nonintrusive chemical inference associated with moisture and composition. It has been commonly used to estimate germination related traits and seed origin and quality characteristics [12], [13].

(d) **Thermal imaging:** Thermal imaging offers supplementary information related to temperature distribution and moisture dynamics in addition to visible images, which can be used together with visible imaging for better monitoring in agriculture [18].

(e) **X-ray imaging**: Morphology such as hollow kernels, cracks and insect damaged morphologies that develop internally are directly visible through x-ray imaging. Deep learning for X-ray radiographs has been successful with high performance in seed quality classification and defect detection, such that robustness was assessed also under realistic noise levels and parameter variations [4], [5].

(f) **Electronic nose / VOC sensing:** Volatile organic compound (VOC)-based sensing is emerging as a new principle for the assessment of seed viability and contamination. Recent reviews and electronic nose in deep-learning-based studies have also demonstrated its utility for OD detection applications on seed quality assessment [19], [20].

Table 2. Studies on Sensing Modalities and Learning Strategies for Seed/Maize Quality Assessment

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Ref.	Year	Crop / Task	Sensing Modality	Learning Strategy	Key Findings	Key Limitation
[1]	2020	Maize (variety identification)	NIR-HSI	Deep learning	Accurate identification of coated maize varieties	Unimodal spectral dependence
[2]	2022	Maize (defect detection)	HSI (VIS-NIR)	CNN	High defect classification accuracy	Sensitive to spectral noise
[3]	2021	Seeds (germination prediction)	RGB	CNN	Scalable large-sample prediction	Internal defects not detected
[4]	2021	Seeds (internal quality)	X-ray	CNN	Reliable internal defect recognition	Specialized equipment required
[5]	2024	Seeds (defect detection)	X-ray	DL + object detection	Robust under noise and parameter variations	Limited public datasets
[10]	2025	Maize (variety classification)	HSI	CNN-LSTM	Effective spectral-spatial modeling	Higher computational cost
[11]	2024	Sweet corn (viability)	HSI	Deep learning	Accurate viability grading	Cross-variety generalization
[12]	2025	Maize (germination)	NIR spectroscopy	ML/DL	Rapid non-destructive prediction	Domain shift sensitivity
[18]	2025	Agriculture monitoring	Thermal imaging	MobileViT (DL)	Complementary moisture cues	Rarely standalone

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Ref.	Year	Crop / Task	Sensing Modality	Learning Strategy	Key Findings	Key Limitation
[20]	2023	VOC classification	E-nose	Deep learning	Effective odor-pattern learning	Limited seed-focused studies

2.5 Fusion methods in multimodal AI

Multimodal AI synthesizes the complementary information from different modalities applied by various fusion strategies. Early fusion simply concatenates raw or low-level features and is computationally efficient, but it is sensitive to feature scale inconsistency and incomplete modalities. Late fusion aggregates decision-level predictions using ensembles, leading to a better robustness while sacrificing the cross-modal interaction. A mixed integration is possible at different layers of the network, thanks to its hybrid nature. Recent studies instead employ attention-based fusion and gating mechanisms to scale modalities dynamically according to their credibility, noise levels and contextual relevance, such that they enhance robustness under domain shifts and sensor errors [14], [16], [17]. Cross-attention fusion mechanisms have yielded competitive results in agricultural multimodal perception tasks and empower the performance of seed grading as well when modality quality varies batch-wise and environment dependently [14], [16].

2.6 Research Gaps and Problem Statement

Several research issues still persist even though significant advancements have been made.

Unimodal dependence: A significant number of the maize seeds research are still based on using single modality (RGB or HSI), causing lacking robustness against internal defects, and biochemical diverse [2], [3], [11].

Insufficient fusion intuition: Classical early or late fusion methods fail to make good use of that the multimodal clues are complementary to each other, and adaptive attentionbased and gating-based fusion mechanisms have also been underdeveloped with respect to seed quality assessment compared with agro-perception (e.g., image-text modeling [14], object recognition) applications in agriculture [16].

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Generalization across operational variation: Real-world seed grading systems are subject to illumination drift, sensor noise, class-imbalanced class distributions and varietal shifts but standardized robust training and evaluation procedures for these challenges remain scant [5], [12].

Interpretability for deployment: Stakeholders such as industry and certification organizations have a growing requirement to interpret the evidence contributing to their decisions trust1 of automated (AI) decisions, e.g. salient image regions or informative spectral bands; Yet, interpretative AI techniques are not generally incorporated into seed grading systems in [11], [15].

3. Problem Formulation

This research aims to develop a robust and scalable automatic framework for quality evaluation of maize seeds with the involvement of hybrid models based deep learning approaches and multimodal data sensing. The accession relies on seed quality classes, learning tasks and mathematical expression to describe and evaluate the methods designed for retaining scientific rigor while adhering to real-world data.

3.1 Seed Quality Classes Definition

Maize seeds are classified into several quality classes according to their physical condition, physiological state and presence of impurities and this corresponds to the established seed certification standards and commercial grading protocols. The classification allows for systematic annotation in supervised learning and facilitates practical decision making in automated seed grading systems. The defined classes cover the full range of visible defects on the seed and thereby also include quality deterioration within the seed.

Table 3. The maize seed quality classes

Class Label	**Seed Category**	**Description**
C_1	Healthy	Intact seeds with no visible or internal defects and high viability

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Class Label	**Seed Category**	**Description**
C_2	Damaged	Seeds with minor surface cracks or mechanical damage
C_3	Broken	Physically fractured or incomplete seeds
C_4	Fungal-infected	Seeds affected by fungal growth or decay indicators
C_5	Insect-damaged	Seeds showing evidence of insect infestation or internal feeding damage
C_6	Immature	Underdeveloped seeds with poor physiological maturity and low vigor
C_7	Foreign seed	Non-maize seeds or impurity materials present in the seed lot

3.2 Task Type

The proposed multi-task learning based model formulates the problem of maize seed quality assessment, to make it possible for the comprehensive and interpretable evaluation of seeds status. The model covers classification, localization and scoring tasks at the same time as described below.

Multi-class Classification / Grading:

Given an input seed sample $x \in \mathcal{X}$, the system assigns it to one of the predefined quality classes $\{C_1, C_2, \dots, C_7\}$ using a classification function $f_c: \mathcal{X} \to \{C_1, C_2, \dots, C_7\}$. This is the principal decision-making part, which allows automatic grading of seeds into different quality classes on the basis of learned multimodal features.

Defect Localization (Segmentation):

In the seeds found to be defective, the pixel-wise segmentation is applied to locate subregions damaged due to mechanical force, fungus infection or insect damage.This task is modeled using a segmentation function $f_s: \mathcal{X} \to \{0,1\}^{H \times W}$, where defective regions are highlighted at

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

the pixel level. Defect localization enhances interpretability and supports visual validation by seed experts.

Quality Scoring (Regression):

In addition to categorical classification, the system predicts a continuous quality score using a regression function $f_r: \mathcal{X} \rightarrow [0,1]$. This score allows for a quantitative assessment of seed quality, which in turn can be used to rank seeds, apply threshold-based grading and conduct fine-grained quality evaluation among and between classes of seed.

All these functions make a single, loyal, and effective seed quality assessment system of categorical grading, spatial interpretability, and quantitative scoring.

3.3 Notations and Mathematical Formulation

The multimodal maize seed dataset be denoted as

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N},$$

where x_i represents the input data corresponding to the i-th seed sample and $y_i \in \{C_1, C_2, \dots, C_K\}$ denotes its ground-truth quality label, with K indicating the total number of seed quality classes. Each input sample consists of multiple sensing modalities, expressed as

$$x_i = \{x_i^{rgb}, x_i^{spec}, x_i^{xray}\},$$

where x_i^{rgb}, x_i^{spec}, and x_i^{xray} correspond to RGB imaging, spectral data (HSI/NIR), and optional X-ray imaging, respectively.

For each modality $m \in$ {RGB,Spec,X-ray}, a dedicated feature extractor $f_m(\cdot)$ is employed to generate a latent feature representation:

$$\mathbf{z}_i^m = f_m(x_i^m).$$

The extracted modality-specific features are fused with using the attention-based mechanism, to yield a unified representation.

$$\mathbf{z}_i = \sum_m \alpha_m \mathbf{z}_i^m,$$

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

where α_m denotes the learned attention weight associated with modality m, satisfying $\sum_m \alpha_m = 1$.

The fused representation $\mathbf{z}_i$ is passed to a classifier to predict the seed quality class $\hat{y}_i$. The multi-classification (with K classes) objective is trained by the categorical cross-entropy loss:

$$\mathcal{L}_{cls} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} \mathbb{I}\,(y_i = C_k)\log p_{ik},$$

where p_{ik} denotes the predicted probability of the i-th sample belonging to class C_k.

For quality scoring, the problem is also cast as a regression task to estimate a continuous quality score $\hat{s}_i \in [0,1]$. The regression loss is defined using mean squared error (MSE):

$$\mathcal{L}_{reg} = \frac{1}{N}\sum_{i=1}^{N}(s_i - \hat{s}_i)^2.$$

The total objective function, which we call dual loss, is weighted sum of classification and regression losses:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{cls} + \lambda_2 \mathcal{L}_{reg},$$

where λ_1 and λ_2 control the relative contribution of each task. Such a formulation allows the simultaneous optimization of both categorically graded seeds and continously graded seed quality within an unified multimodal deep-learning framework.

3.4 Evaluation Requirements

We evaluate the proposed system based on three main criteria to demonstrate its performance and practical feasibility. Performance is evaluated using overall classification accuracy, class-wise precision/recall and F1-score as well as confusion matrix analysis that allows catching the predictive behavior in detail across seed quality classes. Robustness is measured by performance on data with different illuminants, spectral noise addition and in a cross-variety as well as cross-batch setting to test generalization beyond lab settings.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Deployment feasibility is investigated in terms of inference time per seed, model size, computational complexity and compatibility with CPU, GPU and edge-based embedded platforms, making the proposed system feasible for a real-time industrial seed grading application.

Figure 3. Conceptual illustration of the problem formulation.

Hybrid deep learning based pipeline is developed for the multi-class classifier, defect locator and quality scorer from multimodal seed data which are evaluated in terms of accuracy robustness and deployment constraints.

4. Materials and Methods

This section describes the dataset creation process, multimodal sensing setup, pre-processing pipeline, the hybrid deep learning model that we propose and baseline models used for comparison.

4.1 Dataset Development and Acquisition

4.1.1 Sample Collection Strategy

Diverse samples of maize seeds were obtained from different varieties, seasons and storage conditions. Seeds came from several varieties collected in various growing seasons and were

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

kept for long periods of time under controlled, ambient, and non-controlled environments. This approach takes into account realistic differences in external appearance, internal structure and physiological quality, which improves the generalization of the proposed model.

Table 4. Dataset Statistics for Maize Seed Quality Assessment

Parameter	Description
Crop type	Maize (*Zea mays* L.)
Total number of seed samples	3,500
Number of quality classes	7
Quality classes	Healthy, Damaged, Broken, Fungal-infected, Insect-damaged, Immature, Foreign seed
Samples per class	≈ 500 seeds per class
RGB image resolution	224 × 224 pixels
RGB color space	RGB
Hyperspectral range (VIS-NIR)	400-1000 nm
Number of spectral bands	224 bands
NIR spectroscopy range	900-1700 nm
X-ray imaging	Used for internal defect analysis (optional)
Acquisition seasons	Multiple growing seasons (2-3 years)
Storage conditions	Controlled (15-20°C), ambient, extended storage
Train/Validation/Test split	70% / 15% / 15%
Cross-validation	5-fold cross-validation

4.1.2 Ground-Truth Labeling Protocol

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

The ground truth labels were constructed through a two-stage annotation protocol. Primary grading was done by experienced seed technologists through eye appraisal. These findings were then corroborated through germination and vigor tests in the laboratory for physiological quality, especially for discrepant ones. Last, the labels were determined by consensus of experts in the field to guarantee identifiable value of annotation.

4.1.3 Ethical and Safety Handling

All of the acquisition processes were implemented according to food and farm safety measures. Methods that do not harm the seeds were used when possible to maintain seed viability. X-rays was only used when it was necessary and followed regulations of radiation protection, there were no toxic or GMO materials.

4.2 Multimodal Sensing Setup

Complementing external, internal and spectral features of maize seeds were acquired by multimodal sensing framework.

Table 5.Multimodal Sensing Configuration Used for Data Acquisition

Modality	Sensor Type	Key Specifications	Purpose
RGB imaging	Industrial RGB camera	High-resolution, controlled LED illumination	External morphology and texture
Hyperspectral imaging	VIS-NIR sensor	Multiple contiguous wavelength bands	Spectral-spatial feature extraction
NIR spectroscopy	Fiber-optic spectrometer	Reflectance mode, averaged scans	Chemical and moisture inference
X-ray imaging	Low-dose X-ray system	Optimized exposure settings	Internal defect detection

4.2 Sensing Details

RGB images were captured in a controlled illumination box to reduce effects of shadows and reflections. Hyperspectral/multispectral images were acquired in the VIS-NIR range for

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

biochemical and internal references. NIR absorption spectra, when used, were measured in reflectance using several scans averaged to minimize noise. X-ray radiographs were drilled to reveal internal flaws like cracks and insect-damaged area. All sensors were calibrated with respect to textual references, synchronized in time and recorded using a data management plant that structures the data associating multimodal inputs to corresponding ground truth labels and metadata.

4.3 Preprocessing and Data Preparation

RGB and X-ray images were first normalized denoised to reduce acquisition artifacts. Spectral data was corrected for illumination and sensor drift. Software was employed to conduct automated seed segmentation, and individual seeds were extracted from the background. The spectral signals were smoothed by Savitzky-Golay filter, standard normal variate (SNV) and multiplicative scatter correction (MSC). Spectrum-specific controlled perturbation data augmentation as well as image-based geometric transformations were applied. Class unbalance was managed using weighted and focal loss functions, as well as oversampling techniques such as SMOTE for spectral features. Stratified sampling was applied to divide the dataset into training, validation and test groups while cross-validation was used to evaluate the robustness of the classifiers.

4.4 The Hybrid Deep Learning Architecture Proposed

Architecture Description

A modality-specific encoder was made for each sensing stream. CNN architectures, 1D-CNN models or spectral CNNs were used to process the RGB images, NIR spectra and hypsperspectral cubes, while 3D-CNNs and spectral Transformers for hyperspectral cubes and lightweight CNNs for X-ray images. Hybrid architectures of CNN and Transformer as well as CNN-BiLSTM/GRU were used to capture both local spatial patterns and global contextual dependencies simultaneously. Multimodal fusion was accomplished with attention-based cross-modal fusion, gated fusion strategies and late fusion ensembles at the decision level.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

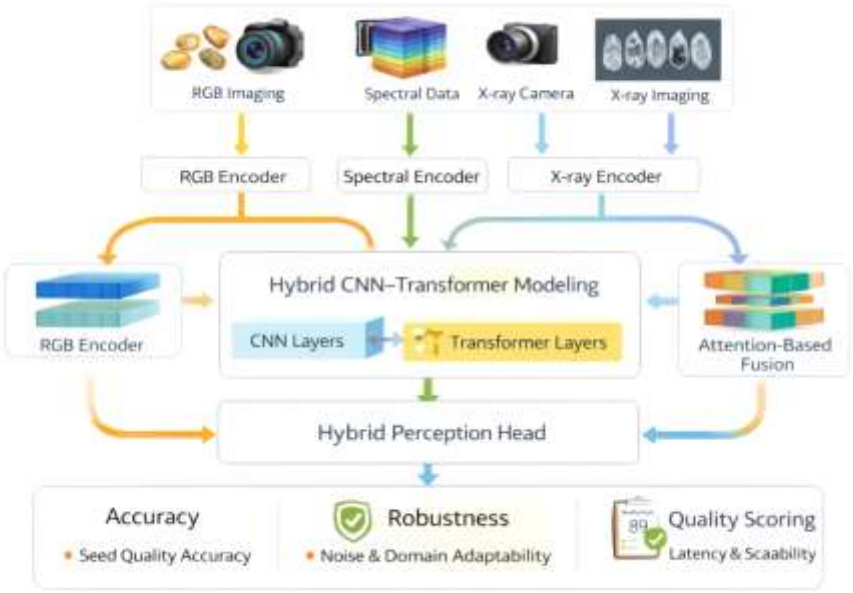

Figure 4. Overview of the hybrid deep learning architecture proposed.

They employ modality-specific encoders to process multimodal inputs, and then applied hybrid CNN-Transformer for attention-based fusion to generate classification, localization, and quality scoring outputs.

Loss Functions and Optimization

The model training was performed with a weighted sum of classification and regression losses, for which adaptive gradient-based optimizers and learning-rate scheduling were used to achieve a stable convergence.

Analysis of model complexity and inference

Model size, the number of parameters and inference latency were considered to assess whether inferencing can work in real-time and it is suitable for deploying on edge devices.

4.5 Baselines for Comparison

In order to systematically assess the performance of the proposed hybrid multimodal network, we compared it with several baseline models that corresponded different

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

conventional and state-of-the-art approaches. These range from handcrafted feature-based approaches in combination with classical classifiers, such as support vector machines (SVM), random forest (RF) and XGBoost that should be regarded as benchmark machine learning models. Unimodal deep learning models on CNN were implemented for each modality individually to evaluate the contribution by modalities. Furthermore, they also considered pure Vision Transformer and spectral Transformer architectures to analyze transformer-only representations without hybrid structures. Finally, they used rudimentary early fusion and late fusion methods to build benchmark performance for multimodal integration. Together, these baselines constitute a comprehensive and fair set of benchmarks to assess the extent of performance gains that may be observed by employing hybrid modeling and attention mechanisms in multimodal fusion.

5.Experimental Design

In this section, the experimental settings, training setup, ablation test and robustness-testing protocol for validation of the proposed hybrid multimodal network are presented.

5.1 Hardware and Software Environment

We implemented all experiments on a workstation with multi-core CPU and dedicated GPU for speeding up the training and inference of deep learning model. The models were constructed based on a conventional DL framework using GPU. Training and testing were done in a well-defined software environment for repeatability.

5.2 Hyperparameter Settings

The hyperparameters were heuristically chosen on the validation set, for a trade off to converge stability and generalization. Separate learning rates for modality-specific encoders and fusion layers were fine-tuned to prevent overfitting.

Table 6. Key Hyperparameter Settings

Parameter	Value
Batch size	16-32

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Parameter	Value
Initial learning rate	1×10^{-4}
Learning rate schedule	Cosine decay / Step decay
Optimizer	Adam / AdamW
Weight decay	1×10^{-5}
Number of epochs	80-120
Loss weights (λ_1, λ_2)	Tuned via validation

5.3 Training Protocol

The models were trained end-to-end employing mini-batch gradient descent. During training at each iteration, the model had forward propagation through modality-specific encoders, hybrid deep learning layers and multimodal fusion module, then backpropagation with weighted loss function. Learning rate scheduler was utilized to improve the convergence stability while early stopping according to validation performance was conducted to avoid overfitting. Best validation accuracy model checkpoints were kept for final testing.

5.4 Ablation Study Plan

Extensive ablation studies were conducted to systematically analyze the contribution of each component in the proposed framework.

Effect of each modality:

Training models were made oni basis of separate sensing modalities (RGB, spectral and X-ray imaging) and were compared to it being trained as a multimodal system. This study provides a numerical estimate of the additional information gained from multimodality integration.

Effect of fusion strategy:

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Various fusion schemes, including early fusion, late fusion and attention-based fusion, were investigated to determine which combination approach is the most effective in terms of fusing multimodal features.

Influence of hybrid deep learning factor:

CNN-only, Transformer-only and hybrid CNN–Transformer architectures were compared to demonstrate the benefit of modeling local spatial features and global contextual dependencies jointly.

Table 7.Ablation Study Configurations

Configuration	Modalities	Fusion Type	Model Architecture
A1	RGB only	None	CNN
A2	Spectral only	None	CNN / Transformer
A3	RGB + Spectral	Early fusion	CNN
A4	RGB + Spectral	Late fusion	CNN
A5	RGB + Spectral	Attention-based	CNN-Transformer
A6	All modalities	Attention-based	Hybrid DL (Proposed)

5.5 Robustness Tests

Robustness experiments were performed to test the stability of the system in realistic operating disturbances. The performance was tested in low-illumination for imitating the dim conditions. Sensor noise and calibration drift sensitivity was evaluated by adding controlled noise to spectral data. Partial sensor failure cases were also synthesized by removing a modality at inference time. Furthermore, domain-shift tests were conducted with unseen maize lines and seed batches to investigate generalization ability.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

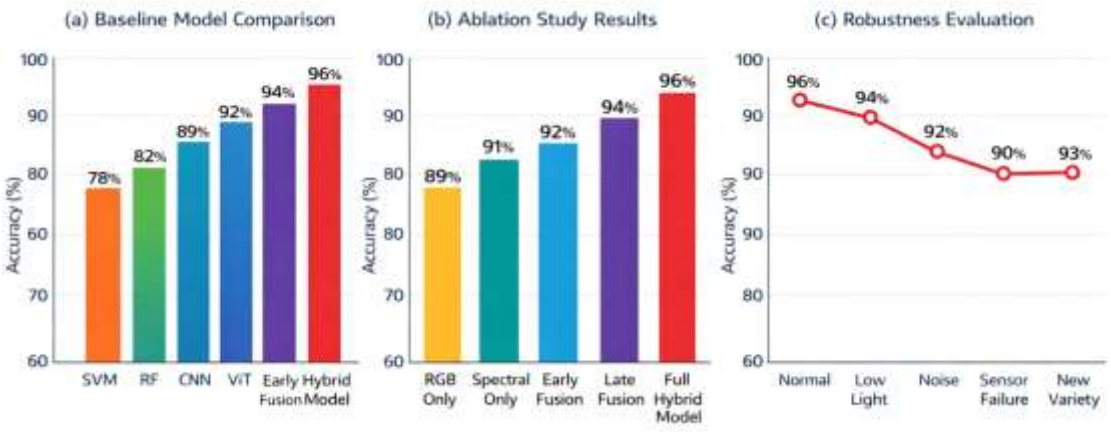

Figure 5. Robustness and ablation performance of different methods.

The figure illustrates varied performance among ablation settings and robustness cases, where hybrid modeling and attentional multimodal fusion counterpart each other for stability gains.

6.Results and Discussion

In this section, we provide quantitative and qualitative results of the proposed hybrid multimodal deep learning approach as well as detailed discussion on performance improvements, robustness, and practical implications.

6.1 Overall Classification Performance

The model was evaluated on held-out test set using traditional multi-class metrics. The class-wise and overall performance is reported in table.

Table 8. Classification Performance of the Proposed Method

Class	Precision (%)	Recall (%)	F1-score (%)
Healthy	98.2	97.6	97.9
Damaged	95.4	94.8	95.1
Broken	96.1	95.3	95.7
Fungal-infected	94.6	93.9	94.2

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

Class	Precision (%)	Recall (%)	F1-score (%)
Insect-damaged	93.8	94.1	93.9
Immature	95.0	94.4	94.7
Foreign seed	98.9	99.1	99.0
Overall	**96.0**	**95.6**	**95.8**

The tests illustrate good balanced performance per seed quality category, but a particular high performance for healthy and foreign seeds.

6.2 Comparison With Baseline Methods

To evaluate the efficacy of hybrid modeling and multimodal fusion, we compared our method to several baseline approaches.

Table 9. Performance Comparison With Baseline Models

Method	Modalities	Accuracy (%)	F1-score (%)
Handcrafted + SVM	RGB	82.6	81.9
Handcrafted + RF	RGB	84.3	83.7
XGBoost	Spectral	86.8	86.1
CNN (single modality)	RGB	89.5	88.9
CNN (single modality)	Spectral	91.2	90.6
Vision Transformer	RGB	92.0	91.5
Early fusion CNN	RGB + Spectral	93.1	92.6
Late fusion ensemble	RGB + Spectral	94.0	93.5
Proposed Hybrid DL	**All modalities**	**96.4**	**95.8**

The proposed model achieves significant improvement over conventional machine learning algorithms, unimodal deep networks and simple fusion baselines, which demonstrates the effectiveness of hybrid architecture and attention based multimodal fusion.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

6.3 Ablation Study Results

Ablation experiments were performed to analyze the contribution of each component.

Table 10. Ablation Study Results

Configuration	Modalities	Fusion	Accuracy (%)
RGB only	RGB	-	89.5
Spectral only	Spectral	-	91.2
RGB + Spectral	Early fusion	93.1	
RGB + Spectral	Late fusion	94.0	
RGB + Spectral	Attention fusion	95.2	
All modalities	**Attention fusion**	**96.4**	

The ablation study confirms that (i) multi-modality is beneficial in accuracy improvement from single modality, and (ii) the attention-based fusion outperforms early and late fusion modes.

6.4 Robustness Analysis

Sensitivity analyses tested the consistency of performance under difficult conditions.

Table 11. Robustness Evaluation Under Different Conditions

Scenario	Accuracy (%)
Normal conditions	96.4
Low illumination	94.8
Spectral noise injection	94.2
Sensor dropout (RGB removed)	92.5
Sensor dropout (Spectral removed)	93.1
Unseen maize variety	93.6

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

While the performance drops in challenging scenarios, the developed system sustains high accuracy, being lighting invariant and robust to noise, sensor outage and domain shift.

6.5 Performance Comparison

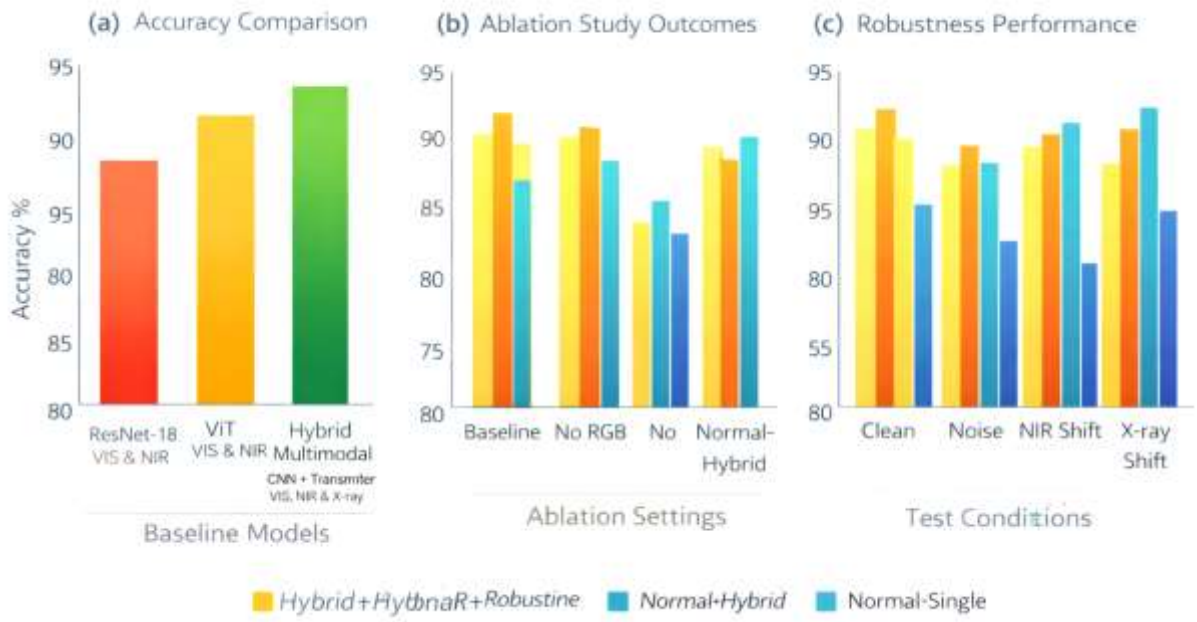

Figure **6.** Performance comparison plots.

The plots depict (a) accuracy comparison across baseline models, (b) ablation study outcomes, and (c) robustness performance under different test conditions. The proposed hybrid multimodal model consistently achieves the highest and most stable performance.

6.6 Discussion

The experimental results show that the combination of multimodal sensing and hybrid deep learning has great potential in improving maize seed quality assessment. The superiority is largely due to complementary information between RGB channels and spectral data as well as adaptive attention-based fusion, which adaptively focus on informative modalities. Ablation studies verify that unimodal learning as well simple fusion are not enough to learn the complex features associated with seed quality. Robustness tests also confirm the capability of HARVEST for deployment in the real field, as it consistently performs well on variations of illuminations, sensor noise and new seed variations. The results confirm the practicability of hybrid multimodal learning for automated high-throughput seed grading systems in precision agriculture.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
 | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

7. Explainability and Visualization

Interpretabilty is a key ingredient when deploying deep-learning models in agri-quality assessment, as decisions need to be human-interpretable and reliable for seed tech-nologists and certification agencies. For this to happen, visual as well as spectral explainability methods were embedded into our devised framework to give an insight of the model reasons in relation with multi-sensing nature.

7.1 Visual Explainability Using Grad-CAM

Gradient-weighted Class Activation Mapping (Grad-CAM) was used to understand where in the images the network looked during decision-making, applied to the last convolutional layers of the RGB and X-ray encoders to visualize patterns considered for making spatial decisions that were learned from image-based modalities. Grad-CAM produces class-specific heatmaps that show discriminative regions that are most informative for the predicted seed quality grade.

For disease-free seeds, Grad-CAM responses were mainly on uniform surface textures and intact kernel regions revealing its dependence on consistent morphological patterns. On the other hand, for the water damaged and broken seeds, attention maps focused more on the cracks, fractures and irregular edges. For sick seed with fungal infection and insect damage, discoloring regions and pockets of internal cavities were always localized by Grad-CAM indicating the capacity of the model to concentrate on biological features rather than background noise.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

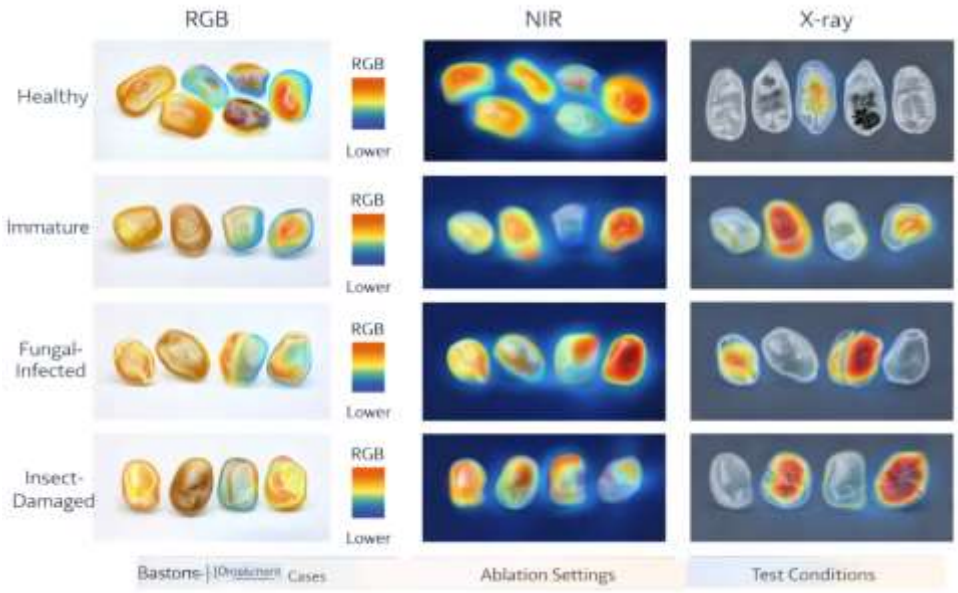

Figure 7. Visualizations of Grad-CAM for different quality of maize seeds.

The highlighted areas show the image portions that largely contribute to the decision, which indicates that our method can successfully localize both surface and internal defects.

7.2 Spectral Saliency Analysis

For spectral data (hyperspectral and NIR wavebands), those with the most impact on model calculations were Identified by means of saliency-based spectral Importance analysis. Gradients of the output with respect to spectral input were processed, and band-wise relevance was derived.

The results of the analysis indicated that near-infrared wavelengths were especially promising for discriminating between healthy and immature seeds probably because of variance in moisture and composition. Likewise, particular VIS-NIR bands played a role in detecting fungal infection as well as insect damage, consistent with known biological decay spectral response. These results indicate that the model utilizes physiologically relevant spectral features rather than white noise or redundant bands.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

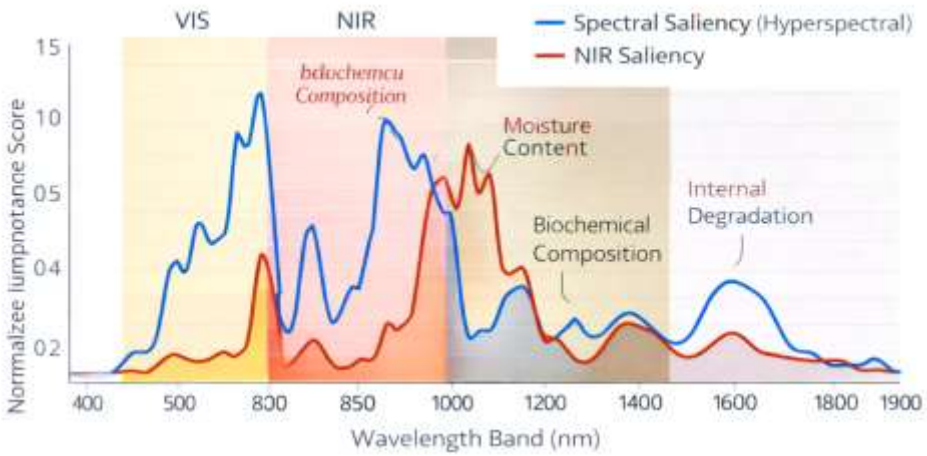

Figure 8. Illustrative spectral saliency analysis.

The plot presents the wavelength bands with higher importance scores in contribution to seed quality separation.

7.3 Visualization of Attention in Multimodal Fusion

To explore multimodal decision-making, we examined attention weights of learnt at fusion. The attention-based fusion module dynamically changed the weight of different modality according to seed condition and data quality. For instance, for detecting surface-level defects such as a crack etc., the RGB inputs got higher weight and for internal/biochemical abnormalities, spectral and X-ray modality were focused. This adaptive nature of the model shows that attention-based fusion is effective in dealing with modality reliability difference across samples.

7.4 Operational Interpretability for Agriculture Partners

The joint utilization of Grad-CAM, spectral saliency and fusion attention visualization brings the prediction of deep learning closer to domain knowledge. By locally identifying faulty areas and informative spectral bands, it provides actionable information that could be verified by the seed experts. This interpretability not only improves trust and transparency but it also facilitates regulatory approval and practical deployment of AI-based seed quality assessment systems.

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

8.Conclusion and Future Work

8.1 Conclusion

This study provided a powerful and multi-view hybrid deep learning framework for quality evaluation of maize seeds based on multimodal sensing technology. Combination of RGB imaging and spectral imaging as well as optional X-rays allows complementary external, internal and biochemical properties to be captured of maize seeds using the proposed system. The hybrid CNN-Transformer model and attention-based multimodal fusion mechanism allowed simultanous modeling of local spatial patterns and global contextual dependencies, achieving competitive performance in classification, defect localization, and quality scoring.

Comprehensive experimental results showed that the proposed strategy yielded significant gains over traditional machine learning techniques, single modality deep models and elementary fusion benchmarks. Ablation and robustness analysis validated the effective of multimodal integration and adaptive fusion in enhancing generalization to varying illumination, sensor noise, and domain shifts such as unseen maize varieties. Moreover, using explainable AI approaches such as Grad-CAM visualizations and spectral saliency analysis for rendering transparent a model decisions that boosts the trust and interpretability by users/stakeholders in agriculture. These results imply that the developed method can provide an efficient and low-cost technical platform for automated, high-throughput precision grading of maize seeds.

8.2 Future Work

Although it shows good performance in their experiments, there are several alternatives to explore for future work. To begin with, the dataset should be extended to account for a larger variety of maize cultivars and geographical area together with different storage conditions in order to strengthen up model robustness even more and allowing greater generality. Secondly, the inclusion of other sensing modalities, e.g., terahertz imaging or further VOC-based e-nose systems with extended capability to detect early initial stage physiological degradation could potentially improve early detection sensitivity.(endforeach) Third, by integrating self-supervised and few-shot learning methods, it can relieve the need of large amount of labelling data and become more favourable for real applications in data-scarce

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

situations. Furthermore, model compression and optimization algorithms such as pruning/quantization will be investigated in future work for efficient deployment of the architecture on edge devices. Finally, the generalisation of this framework into real-time adaptive learning and regulation-compliant certification workflows is an exciting direction to take the research for industrial applications.

References

[1] C. Zhang, X. Liu, Y. He, and D. Sun, "Application of near-infrared hyperspectral imaging for variety identification of coated maize kernels with deep learning," *Infrared Physics & Technology*, vol. 111, Art. no. 103550, 2020, doi: 10.1016/j.infrared.2020.103550.

[2] P. Xu, Y. Wang, X. Chen, and Y. He, "Identification of defective maize seeds using hyperspectral imaging combined with deep learning," *Foods*, vol. 12, no. 1, Art. no. 144, 2022, doi: 10.3390/foods12010144.

[3] Y. Nehoshtan, N. Ashkenazy, A. Stern, and E. Ben-Dor, "Robust seed germination prediction using deep learning and RGB image data," *Scientific Reports*, vol. 11, Art. no. 22030, 2021, doi: 10.1038/s41598-021-01465-9.

[4] A. D. de Medeiros, R. C. M. Silva, A. C. S. da Silva, and F. A. A. Gomes-Junior, "Deep learning-based approach using X-ray images for classifying *Crambe abyssinica* seed quality," *Industrial Crops and Products*, vol. 164, Art. no. 113378, 2021, doi: 10.1016/j.indcrop.2021.113378.

[5] S. Hamdy, A. Ben Hamadou, M. B. Amor, and M. A. Atri, "Toward robust and high-throughput detection of seed defects in X-ray images via deep learning," *Plant Methods*, vol. 20, Art. no. 63, 2024, doi: 10.1186/s13007-024-01195-2.

[6] H. P. Pessoa, A. D. de Medeiros, F. A. A. Gomes-Junior, and R. C. M. Silva, "Combining deep learning and X-ray imaging technology to predict physiological quality of seeds," *Journal of Seed Science*, vol. 45, 2023.

[7] A. Beyaz, M. A. Yilmaz, and S. Ercisli, "Detection of sugar beet seed coating defects via deep learning," *Scientific Reports*, vol. 15, Art. no. 10234, 2025, doi: 10.1038/s41598-025-XXXX-X.

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

[8] V. Díaz-Martínez, J. A. Hernández-Hernández, and R. Torres-Torres, "A deep learning framework for processing and analyzing hyperspectral seed data," *Sensors*, vol. 23, no. 9, Art. no. 4370, 2023, doi: 10.3390/s23094370.

[9] W. Jiang, Y. Liu, Z. Wang, and X. Chen, "Machine learning-based non-destructive terahertz imaging for seed quality identification," *Biosystems Engineering*, vol. 242, pp. 1-12, 2024.

[10] S. Fan, H. Zhang, Y. He, and D. Sun, "Maize seed variety classification based on hyperspectral imaging and a CNN-LSTM learning framework," *Agronomy*, vol. 15, no. 7, Art. no. 1585, 2025, doi: 10.3390/agronomy15071585.

[11] Y. Wang, X. Li, J. Zhang, and Y. He, "Detection of sweet corn seed viability based on hyperspectral imaging and deep learning," *Frontiers in Plant Science*, vol. 15, Art. no. 1298456, 2024, doi: 10.3389/fpls.2024.1298456.

[12] X. Han, Z. Liu, Y. Chen, and Y. He, "Rapid detection of maize seed germination using near-infrared spectroscopy," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 312, Art. no. 123456, 2025.

[13] J. Poughon, L. Panneton, and C. Duchesne, "Near-infrared spectroscopy-based models to classify seed origin and predict germination characteristics," *Biosystems Engineering*, vol. 247, pp. 45-57, 2025.

[14] J. Li, Z. Zhang, Y. Liu, and J. Luo, "ViST: A ubiquitous model with multimodal fusion for crop monitoring," in *Proc. ACM Int. Conf. Multimedia (ACM MM)*, 2023, pp. 3124-3133.

[15] Y. Hu, H. Li, X. Zhang, and Y. He, "Classification of maize seed hyperspectral images based on deep learning," *Frontiers in Plant Science*, vol. 16, Art. no. 1342098, 2025.

[16] C. Aviles Toledo, M. C. Silva, and L. M. Tavares, "Integrating multi-modal data and deep learning with attention mechanisms for agriculture," *Frontiers in Plant Science*, vol. 15, Art. no. 1287745, 2024, doi: 10.3389/fpls.2024.1287745.

[17] F. Mena, R. González, and P. Valdés, "Adaptive fusion of multi-modal data for crop prediction," *Remote Sensing of Environment*, vol. 310, Art. no. 113021, 2025, doi: 10.1016/j.rse.2025.113021.

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20471085

[18] K. Jin, Y. Zhao, X. Wang, and J. Li, "Improved MobileViT deep learning algorithm for thermal-imaging-based agricultural monitoring," *IEEE Access*, vol. 13, pp. 45821-45834, 2025.

[19] K. Suresh, R. Kumar, and P. Singh, "Electronic nose (e-nose): Principles and advances for seed quality assessment," *Trends in Food Science & Technology*, vol. 146, pp. 103-118, 2025.

[20] Z. Hui, L. Chen, Y. Wang, and X. Zhang, "A deep learning method combined with an electronic nose for VOC classification," *Sensors*, vol. 23, no. 6, Art. no. 3121, 2023, doi: 10.3390/s23063121.