World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Review Article)

Check for updates

Seed detection and classification with artificial intelligence: A review

Taminul Islam *

School of Computing, Southern Illinois University Carbondale, Carbondale, IL 62901, USA.

World Journal of Advanced Research and Reviews, 2024, 23(03), 508–522

Publication history: Received on 25 July 2024; revised on 31 August 2024; accepted on 03 September 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.23.3.2689

Abstract

Seed detection and classification play a crucial role in the agriculture domain, and artificial intelligence has been increasingly combined with agriculture in various sectors. Manual seed detection and classification are time-consuming and less accurate compared to automated methods. Several research works have been conducted on automatic seed detection and classification using machine learning and deep learning algorithms. This review examined ten experimental research works focusing on seed detection and classification using these algorithms. The approaches, contributions, datasets, data preprocessing, algorithms, results, and limitations of each work were reviewed and presented. The survey revealed that Convolutional Neural Networks (CNN) are the most frequently chosen algorithms for seed classification, with 93% of the reviewed works using CNN for comparing and evaluating their models. Based on the in-depth survey, four recommendations are made for consideration in future experimental analyses of seed detection and classification. These findings highlight the importance of artificial intelligence in advancing seed detection and classification techniques in the agriculture domain.

Keywords: Seed classification survey; Seed detection review; Review on seed identification; Seed analysis; Review in agriculture

1. Introduction

A seed is the fundamental and crucial component for sustainable agriculture. Moreover, the response of all additional factors in the realm of agriculture is heavily contingent upon the caliber of seeds. Accurate identification and categorization of seeds are essential for various purposes, including quality control, inventory management, and research [1]. However, traditional manual methods for seed detection and classification are time-consuming, labor-intensive, and prone to human error. These limitations can lead to inconsistencies, reduced efficiency, and potential economic losses in the agricultural sector.

The advent of artificial intelligence (AI) has opened up new possibilities for automating seed detection and classification processes. AI techniques, particularly machine learning and deep learning algorithms, have the potential to overcome the drawbacks of manual methods and provide more accurate, efficient, and reliable solutions [2] [3]. By leveraging the power of AI, researchers and practitioners in the agriculture domain can develop automated systems that can handle large volumes of seed data, perform high-throughput processing, and ensure consistency and accuracy in seed identification and categorization.

* Corresponding author: Taminul Islam

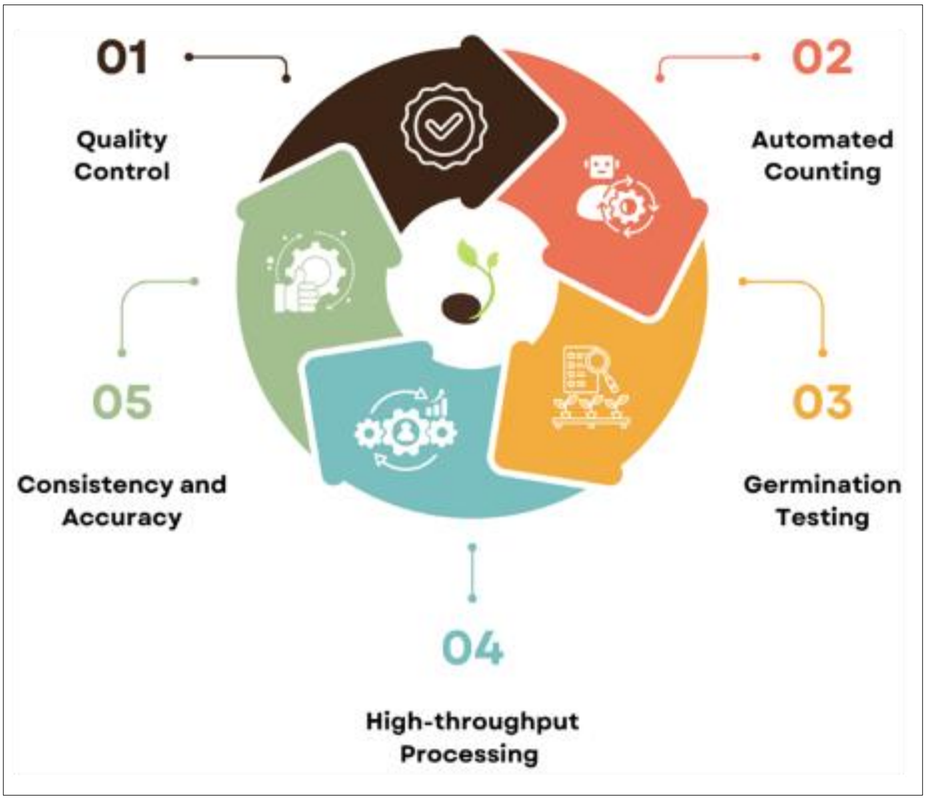

Figure 1 Visualization of the benefits of automated seed detection and classification

The application of AI in seed detection and classification offers several benefits. Firstly, it enables quality control by automatically identifying and separating seeds based on their physical characteristics, such as size, shape, and color [4]. This ensures that only high-quality seeds are selected for planting, leading to improved crop yield and quality. Secondly, AI-based systems can perform automated counting of seeds, which is crucial for inventory management and production planning [5]. Thirdly, AI can assist in germination testing by analyzing seed images and predicting their viability, saving time and resources compared to manual testing methods [6]. Furthermore, AI-powered seed detection and classification systems can handle high-throughput processing, allowing for the analysis of large quantities of seeds in a shorter time frame. This is particularly beneficial for seed companies, research institutions, and agricultural organizations dealing with vast amounts of seed data.

This review aims to provide a comprehensive overview of the current state of research in seed detection and classification using AI techniques. By examining ten experimental research works, this review will shed light on the various approaches, contributions, datasets, data preprocessing methods, algorithms, results, and limitations of each work. The insights gained from this survey will contribute to the advancement of the agriculture domain by highlighting the potential of AI in automating seed detection and classification processes. It will also identify areas for future research and development, guiding researchers and practitioners in their endeavors to improve seed identification and categorization techniques.

The expected outcome of this survey is to provide a valuable resource for researchers, agricultural professionals, and decision-makers in the agriculture domain. By presenting a systematic review of the existing literature, this survey will facilitate a better understanding of the current landscape of AI-based seed detection and classification. It will also offer recommendations for future experimental analyses, empowering researchers to build upon the existing knowledge and develop more robust and efficient AI models for seed identification and categorization. Ultimately, this review aims to contribute to the development of advanced AI technologies that can revolutionize seed detection and classification processes, leading to improved efficiency, accuracy, and productivity in the agriculture sector.

2. Motivation

The motivation for this research stems from the critical importance of seed detection and classification in the agriculture domain and the need to address the limitations of manual methods. Accurate and efficient seed identification and categorization are crucial for ensuring the quality and productivity of crops, as well as for supporting various agricultural processes [7] such as inventory management, research, and seed production. As critical components of agricultural management, seed detection and classification impact seed quality, crop yield, and resource utilization. Manual methods for these tasks are not only labor-intensive but also prone to errors, making them less viable given the scale and diversity of modern agricultural demands [8]. The advent of advanced artificial intelligence technologies, especially in the fields of machine learning and deep learning, offers promising solutions to these challenges [9]. These technologies can automate seed detection and classification processes, providing faster, more accurate, and scalable alternatives to traditional methods [5]. Furthermore, there is a significant need for a systematic review of the existing literature to consolidate knowledge, identify gaps, and suggest future research directions. Such a comprehensive analysis is essential for guiding researchers and practitioners in understanding the current capabilities and limitations of AI in seed classification. Ultimately, this review aims to illuminate how AI-driven approaches can revolutionize seed classification, thereby benefiting stakeholders across the seed industry—from farmers to agricultural technologists—through enhanced decision-making and improved agricultural outcomes.

3. Working Mechanism

This survey aims to provide a comprehensive review of ten experimental research papers focusing on seed detection and classification using artificial intelligence techniques. The primary objective is to analyze and compare the approaches, contributions, datasets, data pre-processing methods, models, results, and limitations presented in each study. To achieve this, a systematic approach was employed to identify and select relevant research papers. The selection process focused on experimental studies that specifically addressed seed detection and classification using AI algorithms (machine learning, deep learning, and computer vision). The selected papers were chosen based on their relevance, quality, and potential impact on the field of agriculture.

Table 1 provides an overview of the titles of the ten papers considered in this survey. These papers represent a diverse range of research efforts in the domain of seed detection and classification, showcasing different techniques, datasets, and evaluation metrics. The publications span from 2020 to 2023.

Table 1 Title, Year, Author, and Publisher's information of ten selected papers

SL.	Author	Title	Year	Publisher
1	Fu et al. [10]	Cultivars identification of oat (Avena sativa L.) seed via multispectral imaging analysis.	2023	Frontiers
2	Gulzar et al. [11]	A Convolution Neural Network-Based Seed Classification System.	2020	MDPI
3	He et al. [12]	Multi-Modal Late Fusion Rice Seed Variety Classification Based on an Improved Voting Method.	2023	MDPI
4	Kiratiratanapruk et al. [13]	Development of Paddy Rice Seed Classification Process using Machine Learning Tech-niques for Automatic Grading Machine.	2020	Hindawi
5	Loddo et al. [14]	A Novel Deep learning based approach for seed image classification and retrieval.	2021	Elsevier
6	Rathnayake et al. [15]	Age Classification of Rice Seeds in Japan Using Gradient-Boosting and ANFIS Algorithms.	2023	MDPI
7	Xu et al. [16]	Research on Maize Seed Classification and Recognition Based on Machine Vision and Deep Learning.	2022	MDPI
8	Peng et al. [17]	Automatic monitoring system for seed germination test based on deep learning	2022	Wiley
9	Zhao et al. [18]	Deep-learning-based automatic evaluation of rice seed germination rate.	2022	Wiley
10	Ouf et al. [19]	Leguminous seeds detection based on convolutional neural networks: Comparison of Faster R-CNN and YOLOv4 on a small custom dataset.	2023	Elsevier

The review process involved a thorough examination of each paper, focusing on key aspects that contribute to the understanding and advancement of AI-based seed detection and classification methods. The approaches section of the review explored the underlying methodologies and algorithms employed in each study. This included an analysis of the specific AI techniques used, such as machine learning algorithms and deep learning architectures. The datasets section provided an overview of the data sources and characteristics used in each study. This included information about the type of seeds, the number of samples, the imaging modalities, and any specific challenges or limitations associated with the datasets. The data pre-processing section delved into the techniques employed to prepare the raw seed data for analysis. This encompassed tasks such as image segmentation, feature extraction, data augmentation, and normalization. This review examined how each study handled data pre-processing and its impact on the overall performance of the seed detection and classification models. The models section focused on the specific AI architectures and configurations used in each research paper. This included an analysis of the network structures, hyperparameters, and training procedures employed. The results part presented the key findings and performance metrics reported in each study, and compared the results across different studies to identify trends and potential areas for improvement. Finally, the limitations section discussed the challenges and constraints encountered in each research paper. These included issues related to dataset size, data quality, computational resources, or the generalizability of the proposed methods to different seed varieties or environmental conditions. This survey also considered potential future directions and recommendations for addressing these limitations.

By conducting a comprehensive review of these ten experimental research papers, this survey aims to provide a holistic understanding of the current state-of-the-art in AI-based seed detection and classification. After reviewing these studies, this survey can summarize the AI models employed in the referenced papers and recommend the most effective algorithms for future research. Additionally, it advises researchers to consider four key aspects in their upcoming work on seed detection and classification.

4. Approaches and Contributions

This section provides a detailed examination of the approaches and contributions of the ten selected research works on AI-based seed detection and classification. The studies surveyed employ a variety of methodologies, ranging from traditional machine learning techniques to advanced deep learning architectures, each tailored to address specific challenges in seed identification and classification. A predominant trend observed across these works is the reliance on Convolutional Neural Networks (CNNs) due to their robust performance in image-based tasks. Furthermore, several studies incorporate innovative data preprocessing techniques, such as multispectral imaging and feature fusion, to enhance model accuracy and generalizability. These contributions not only demonstrate the evolving capabilities of AI in agriculture but also offer valuable insights for optimizing future research efforts in seed detection and classification.

Fu et al. [10] developed a novel approach for rapid and non-destructive identification of oat cultivars by integrating multispectral imaging technology with multivariate data analysis. Their methodology involved capturing high-resolution images of oat seeds using a multispectral imaging system that recorded data at 19 specific wavelengths spanning [20] the visible and near-infrared regions of the electromagnetic spectrum.

From the multispectral images, they extracted a comprehensive set of morphological features for each individual seed, including size, shape, and color parameters. In addition, they obtained detailed spectral reflectance data across the 19 wavelengths, which provided valuable information about the chemical composition and internal structure of the seeds. By combining the morphological and spectral data, they aimed to develop robust models for discriminating between different oat cultivars based on the intrinsic properties of their seeds. This innovative approach leverages the power of imaging technology to capture a wealth of information about seed characteristics, which can then be analyzed using advanced statistical techniques to enable accurate cultivar identification. Overall, this work represents a significant advancement in the field of oat cultivar identification, as it offers a rapid, non-destructive, and potentially high-throughput alternative to traditional methods that rely on time-consuming manual inspection or destructive molecular marker analysis.

Gulzar et al. [11] proposed an efficient model for seed identification and classification based on convolutional neural networks (CNN) [21] and the application of symmetry. The model was trained using transfer learning, focusing more on validation and testing rather than training from scratch. The authors made several key contributions in this work. They conducted a detailed review of the most notable research in the area of seed classification using machine learning and deep learning. They re-introduced the problem of seed classification using a pre-trained VGG16 CNN architecture, classifying 14 different types of seeds, unlike previous models that focused on only one type of seed. He et al. [12] presented a rice variety classification method that performs late fusion of 2D and 3D modalities using an improved voting approach. They collected 2D images and 3D point cloud data of eight common rice varieties using a Raytrix light field camera [22]. The authors developed an improved late-fusion method that generates a dynamically changing scoring vector based on each model's performance. This vector is used to adjust the influence of individual predictions on the final fused result. The collected data was preprocessed and input into modality-specific models to obtain predicted probabilities. The scoring vector is then used to calculate a weighted probability from the different models, with the highest probability class being selected as the final predicted variety. The key contribution is the improved voting method for late fusion, which the authors state is more robust compared to other multimodal fusion approaches, as it is less impacted by single poorly performing models and avoids interference from excessive model homogeneity or suboptimal individual model performance.

Kiratiratanapruk et al. [13] presented a study that utilized machine vision technology to classify 14 Oryza sativa rice varieties. The authors developed a rice varieties classification process consisting of object orientation to align seed images, image screening for outlier/irregular/abnormal or tilted seeds, feature extraction for retrieving physical seed properties, and rice varieties classification using machine learning techniques. The authors' key contributions include collecting a large dataset of over 3,500 seed samples for each of the 14 rice varieties, totaling close to 50,000 images, to cover the diversity within each rice species. They also developed preprocessing methods for seed orientation and quality screening to prepare high-quality input data for the classification models. Furthermore, they compared the performance of traditional machine learning methods with deep learning techniques using pretrained models for rice variety classification. Loddo et al. [14] presented a novel deep learning approach for seed image classification and retrieval. They proposed SeedNet, a new CNN architecture, and compared its performance with several state-of-the-art CNNs on two seed datasets. The authors also compared the deep learning methods with traditional machine learning approaches using handcrafted features. Additionally, they introduced two new preprocessed seed datasets containing single seed images. Their contributions include proposing the accurate and efficient SeedNet CNN, extensively comparing it with ten existing CNNs, comparing CNNs with four classical machine learning techniques, and creating two new seed image datasets through preprocessing. The study aimed to find the best performing model architecture and training options for seed classification and retrieval tasks.

Rathnayake et al. [15] presented a novel machine learning approach for classifying Japanese rice seeds based on their variety and harvest age. They developed a new rice seed dataset containing six rice varieties with three age variations, which is the only dataset labeled by harvested age to the authors' knowledge. The authors investigated various features extracted from RGB images using six feature descriptors to accurately classify the seeds. They proposed a novel machine learning model combining gradient-boosting algorithms [23] with a Cascaded-ANFIS structure for cost-effective and efficient identification of rice seed variety and age. The proposed algorithm's performance was compared with several other feature-based machine learning algorithms. This study aims to provide a solution for replacing complex, power-consuming black box algorithms with a more efficient and effective method for identifying seed variety and age. Xu et al. [16] presented a non-destructive method for automatic identification and classification of different varieties of maize seeds from images using machine vision combined with deep learning. They applied a CNN architecture called P-ResNet for varietal detection. The authors established a seed dataset and divided it into training and validation sets in an 8:2 ratio for experiments. They evaluated and compared the classification performance of various CNN models and used visualization techniques to validate the results. Xu et al. hypothesized that transfer learning can help save model training time by acquiring knowledge learned in other settings for similar tasks, and that CNN models can automatically extract more depth features from images compared to manual feature extraction methods, thus improving classification performance. The main contributions of this work include proposing the P-ResNet architecture for maize seed classification, demonstrating the effectiveness of transfer learning and data augmentation for this task.

Peng et al. [17] present an innovative automatic monitoring system for seed germination tests, addressing the limitations of traditional methods that rely on manual labor and expensive equipment. Their system combines hardware modifications to germination thermostats with a multifunctional software system and a novel deep learning algorithm, DDST-CenterNet, designed for dense small target detection. The authors highlight the system's cost-effectiveness, versatility, and independence from environmental factors such as seed background and lighting conditions. The DDST-CenterNet algorithm demonstrates high accuracy and stability in detecting and classifying seeds, even as seed density increases. Notably, the algorithm's computational efficiency allows for real-time detection at frame rates of at least 10fps. This work represents a significant advancement in seed germination testing, offering a practical and scalable solution that could greatly improve breeding efficiency and reduce reliance on specialized technicians. The system's potential for widespread adoption in agricultural research and industry is evident. Zhao et al. [18] presented a deep learning-based method called YOLO-r for automatically evaluating the germination rate of rice seeds. They developed YOLO-r by improving the YOLOv5 object detection model to better identify small, densely distributed rice seeds. The authors made several contributions to enhance the performance of YOLOv5. They divided the input images into overlapping sub-images to maintain resolution while ensuring efficient processing. The Transformer encoder was incorporated into the backbone and head networks to accurately predict densely distributed small objects using the self-attention mechanism. Additionally, a small target detection layer was added to improve the detection of small-sized rice seeds. Finally, the authors adopted the CDIoU loss function to accelerate the convergence of the predicted bounding box loss and improve network performance.

Ouf [19] presented a comparative study of two deep learning-based models, Faster R-CNN and YOLOv4, for the detection and identification of 11 different types of leguminous seeds. The author manually collected and annotated a diverse dataset of 828 seed images, considering variations in background, crowdedness, shooting angle, and seed combinations. Transfer learning was employed to optimize the models, with Inceptionv2 and CSPDarknet53 serving as the backbone feature extractors for Faster R-CNN and YOLOv4, respectively. The study aimed to address the challenges of detecting small, similarly colored seeds in complex scenarios, with a focus on improving accuracy and runtime. Ouf's work contributes to the field of agricultural object detection by providing a comprehensive evaluation of two state-of-the-art models on a custom leguminous seed dataset, highlighting the potential of YOLOv4 for accurate and efficient seed detection in real-world applications.

5. Dataset

The selection and preparation of appropriate datasets are crucial for developing and evaluating effective seed classification models. This section provides an overview of the datasets used in recent studies on seed classification, highlighting their composition, acquisition methods, and preparation techniques. Researchers have employed a wide variety of datasets, ranging from single-species collections to diverse multi-species compilations. These datasets typically include high-quality images of seeds, often captured under controlled conditions to ensure consistency. Some studies have also incorporated 3D data or images from multiple angles to enhance the robustness of their models.

Fu et al. [10] used a dataset consisting of seeds from 16 oat (Avena sativa L.) cultivars, namely Blade, Deon, Jerry, Kona, Longyan1, Longyan2, Longyan3, Longyan4, Brave1, Morgan, Monica, Tanke, Youmu1, Baiyan7, Dingyan2, and Quebec. For each cultivar, they collected 200 seeds, which were carefully selected to be free from any visible defects or diseases. The seeds of Dingyan2, Baiyan7, and Quebec were sourced from the Academy of Agricultural Science, Dingxi Gansu Province, while the remaining cultivars were obtained from Beijing Best Grass Industry Co., Ltd. [24] To ensure the robustness and reliability of their models, Fu et al. randomly divided the seeds of each cultivar into a training set (140 seeds, 70% of the total samples) and an independent testing set (60 seeds, 30% of the total samples).

Gulzar et al. [11] prepared a dataset consisting of images from 14 different types of commonly known seeds: black-eyed pea, black pepper, chickpea, coriander, corn, cumin, fennel, fenugreek, flax, kidney beans, mustard, onion, pumpkin, and sunflower. Around 100 grams of each seed type was used for experimentation. A smartphone with a 12-megapixel camera was used to capture the digital images of the seeds from a distance of one foot during the day to avoid texture changes and shadow effects. A white ring light and a tripod mobile stand were used to ensure consistent lighting conditions. Approximately 200 images were captured for each seed type to maintain a balanced dataset and prevent biased results. The captured images were then rescaled to 224 × 224 × 3, labeled, and arranged in separate folders to prepare the dataset for training the CNN model. In [12] study, He et al. used a dataset consisting of eight common rice varieties in China: Nanjing9108, Zhenghan10, Hannuo35, Yuanhan35, Huanghuazhan, Hyou518, Liannuo13, and Liusha. For each variety, they collected both 2D images and 3D point cloud data using a Raytrix light field camera [22]. The dataset was carefully prepared by manually screening and cleaning the rice seeds to avoid irregularities such as attached impurities and gaps that could affect the classification results. The selected samples were stored in a dry, low-temperature, and airtight environment to prevent any influence from external factors. In total, the dataset comprised 3,194 samples, with each seed having a corresponding 2D picture and 3D point cloud of both front and back sides. The dataset was then divided into a training set and a test set at a ratio of 8:2, resulting in 2,560 samples in the training set and 634 samples in the test set.

Kiratiratanapruk et al. [13] utilized a private dataset consisting of 14 popular and economically potential Thai rice varieties provided by the Thailand Rice Department. The rice varieties were categorized into three groups based on their planting areas to analyze ambiguity among varieties. The first group (GrpI) contained 7 varieties: CNT1, KDML105, PTT, RD15, RD33, RD51, and RD6. The second group (GrpII) included 7 varieties: PSL2, RD31, RD41, RD47, RD49, RD57, and SPR1. The third group (GrpAll) was a combination of the first two groups, comprising all 14 varieties. Each variety had more than 3,500 seed samples photographed using a flatbed scanner with a special box tray at 600 DPI resolution, resulting in a total of nearly 50,000 images for the entire dataset. The seed samples were collected from various provinces in Thailand to cover different characteristics influenced by diverse producing environments and areas. Loddo et al. [14] utilized two datasets in their research: the Canadian dataset and a local dataset. The Canadian dataset, publicly available and (last updated in February 2019), contains 587 seed images of different sizes, organized into families belonging to the Magnoliophyta phylum. For their experiments, the authors selected 215 images from the six most represented families: Amaranthaceae, Apiaceae, Asteraceae, Brassicaceae, Plantaginaceae, and Solanaceae. The local dataset consists of 3,386 seed samples from 120 plant species of the Fabaceae family, obtained from the Germplasm Bank of Sardinia (BG-SAR), University of Cagliari, Italy. The authors selected 1,988 seeds from the 23 most numerous species for their experiments. Both datasets underwent a preprocessing step to create single seed images, which involved removing scale indicators and creating binary seed masks for the Canadian dataset, and utilizing the blue background for automatic thresholding in the local dataset.

Rathnayake et al. [15] constructed a novel rice seed dataset. The dataset consists of six Japanese rice varieties: Akitakomachi, Fusaotome, Hatsuboshi, Koshihikari, Okiniiri, and Yang DAO-8. Each variety includes samples from three different harvest years (2012, 2016, and 2020), except for Okiniiri, which has samples from 2012 and 2016, and Yang DAO-8, which has samples from 2012 and 2020. The dataset was created using a conveyor belt system that automatically acquires seed images with a smartphone camera equipped with a macro lens, allowing for the capture of detailed surface features and real-time flexibility. The images were preprocessed by removing backgrounds and segmenting individual seeds. In total, the dataset consists of 16 classes, with the number of samples per class ranging from 261 to 509. The complete dataset, which was divided into training and testing sets at a ratio of 7:3, is publicly available on the Kaggle.

Xu et al. [16] used a dataset consisting of 8080 maize seeds from five common varieties in China, namely BaoQiu, KouXian, LiaoGe, ShanCu, and XinNuo. The seeds were provided by the National Seed Breeding Base in Hainan, China [25] , and were selected and certified by experts. The dataset was manually cleaned to remove impurities and dust. The number of seeds varied for each variety due to the influence of seed storage conditions, with 1710 BaoQiu, 1800 KouXian, 570 LiaoGe, 2000 ShanCu, and 2000 XinNuo seeds. The seeds were placed individually on a black background for image acquisition. All images were captured under the same environment, with a camera resolution of 3384 × 2708 pixels. The acquired images were then segmented into individual seed images of 350 × 350 pixels and saved in PNG format. The entire dataset was randomly divided into a training set (80%, 6464 seeds) and a validation set (20%, 1616 seeds) for the experiments.

Zhao et al. [18] collected a total of 21,429 rice seeds to construct the dataset for their research. The seeds were obtained from nine populations with different phenotypic characteristics, varying considerably in length, shape, and color. The collected seeds were divided into two datasets: RiceSeed1 and RiceSeed2. RiceSeed1, consisting of 600 images, was primarily used for training the YOLO-r model. This dataset was split into a training set and a test set in an 8:2 ratio, with the training set containing 480 images and the test set containing 120 images. Each image in RiceSeed1 was manually labeled using the labelImg tool [26], with germinated seeds marked as 'yes' and ungerminated seeds marked as 'no'. RiceSeed2, comprising 1,211 images, was used to evaluate the performance of the trained YOLO-r model. The images in both datasets contained seeds of different sizes, shapes, and colors, as well as impurities such as branch stalks, broken leaves, and rice awns, with the seeds randomly distributed within the images.

Ouf [19] manually collected and annotated a diverse dataset of leguminous seed images for this study. The dataset consists of 828 images, capturing 11 different types of leguminous seeds: Glycine max, Lens culinaris-dark, Lens culinaris-yellow, Lupinus albus, Medicago sativa, Phaseolus vulgaris-pink, Phaseolus vulgaris-red, Phaseolus vulgaris-white, Trifolium alexandrinum, Trigonella foenun graecum, and Vicia faba. The images were taken using two different devices (a SAMSUNG smartphone camera and a Canon digital camera) and feature variations in background (white A4 paper, black pad, dark blue pad, dark green pad, and green pad), seed crowdedness (1 to 50 seeds per image), shooting angles, and heights. The dataset also includes different combinations and arrangements of the 11 seed types. In total, the 828 images contain 9,801 seed objects (labels). The dataset was randomly split into three subsets: train (80%), validation (10%), and test (10%) for model development and evaluation.

Table 2 Summary of Datasets Used of these 10 research

Author	Dataset Description	Samples	Details
Fu et al.	Dataset of seeds from 16 oat cultivars, divided into training and testing sets.	3200 seeds	Training (70%), Testing (30%)
Gulzar et al.	Images of 14 different seed types captured under consistent lighting conditions.	2800 images	Training (80%), Testing (20%)
He et al.	Dataset of 8 common rice varieties in China, including 2D and 3D data.	3194 samples	Training (80%), Testing (20%)

Kiratiratanapruk et al.	Images of 14 Thai rice varieties from various planting areas.	50,000 images	Training (70%), Testing (30%)
Loddo et al.	Two datasets: Canadian and local from BG-SAR, involving multiple seed families.	5581 images	Not specified
Rathnayake et al.	Novel dataset of Japanese rice varieties with samples from different years.	5282 images	Training (70%), Testing (30%)
Xu et al.	Dataset of 8080 maize seeds from five varieties in China.	8080 seeds	Training (80%), Validation (20%)
Yasam et al.	Images of three seed classes, each with a sequence of images.	23,797 images	Not specified
Zhao et al.	Two datasets of rice seeds with detailed classifications and varying conditions.	21,429 seeds	Training (80%), Testing (20%) for RiceSeed1
Ouf et al.	Diverse dataset of leguminous seed images, manually collected and annotated.	828 images	Train (80%), Validation (10%), Test (10%)

6. Data Pre-Processing

Data preprocessing is a crucial step in computer vision and machine learning tasks involving image data, as it prepares the raw data for effective analysis and modeling. In the context of seed classification and germination studies, various researchers have employed diverse preprocessing techniques to enhance the quality of the input images, extract relevant features, and ensure the optimal performance of their models. The following section outlines the preprocessing approaches adopted by these ten studies, highlighting the importance of data preparation for accurate and reliable results.

Fu et al. [10] first used a normalized canonical discriminant analysis (nCDA) in the VideometerLab software to segment the acquired multispectral images and create image masks that isolated the oat seeds from the background and non-seed pixels. This step ensured that only the relevant seed pixels were considered for further analysis. Subsequently, they extracted the morphological features (such as area, length, width, shape, and color parameters) and spectral reflectance data for each individual seed from the pre-processed images.

Gulzar et al. [11] applied several data pre-processing techniques in their work. They used the image augmentation technique to artificially produce more training images using the ImageDataGenerator API from the Keras library. The augmented images were generated by randomly rotating, zooming, adjusting height, and shifting width of the original images. This process helped expose the system to a wide variety of variations in the dataset. The augmented images were then rescaled to a consistent size of 224 × 224 × 3 to ensure compatibility with the proposed model. Additionally, the authors employed a dropout technique to minimize overfitting and improve the model's validation performance.

In [12] work, the dataset was carefully prepared by manually screening and cleaning the rice seeds to avoid irregularities such as attached impurities and gaps that could affect the classification results. The selected samples were stored in a dry, low-temperature, and airtight environment to prevent any influence from external factors. In total, the dataset comprised 3,194 samples, with each seed having a corresponding 2D picture and 3D point cloud of both front and back sides. The dataset was then divided into a training set and a test set at a ratio of 8:2, resulting in 2,560 samples in the training set and 634 samples in the test set.

Kiratiratanapruk et al. [13] employed a preprocessing method consisting of two main steps: (1) seed orientation and (2) seed quality screening. In the seed orientation step, the authors developed a method to examine and rotate the seed body into the horizontal axis direction, ensuring that all seeds' head-and-tail directions were aligned uniformly. This alignment simplified the feature extraction and data analysis processes. The seed quality screening step involved detecting and discarding outlier seeds (seeds with different shapes from the standard) and tilted seeds (seeds not properly aligned horizontally during scanning). The authors utilized the DBSCAN clustering technique to identify outlier seeds and the SVM technique with shape features to classify tilted seeds. These preprocessing steps aimed to ensure high-quality input data for the subsequent classification models.

Loddo et al. [14] applied preprocessing steps to create single seed images for both datasets. For the Canadian dataset, they first removed the scale indicators present in the images. Then, they created binary seed masks by identifying the background's intensity values using the image histogram and excluding them based on a 5% range. The masks were used to identify bounding boxes for each seed, and the main seed image was extracted by replacing non-mask pixel values with the background pixel value. For the local dataset, the authors utilized the blue background to isolate the seeds using automatic thresholding on the blue channel of the RGB images. As the seeds were well-spaced during acquisition, the bounding box of each region was used to create a single image for every seed. Some species were discarded due to low sharpness and small seed size. Xu et al. [16] employed several data pre-processing techniques in their work. First, the acquired high-resolution images containing multiple seeds were segmented into individual seed images of 350 × 350 pixels. Next, data augmentation was applied to the training images, which were randomly rotated, flipped horizontally and vertically, and normalized to extend the dataset and improve the model's classification precision and robustness. Lastly, the entire dataset was randomly divided into a training set (80%, 6464 seeds) and a validation set (20%, 1616 seeds) for the experiments. These pre-processing steps were crucial for preparing the dataset to be used effectively in the convolutional neural network (CNN) models for maize seed classification.

Zhao et al. [18] applied several preprocessing techniques to the images in the dataset. First, they adjusted the contrast and brightness of the images using gamma correction and log transformation. Second, Gaussian noise was added to the RGB channels of the images to simulate the noise introduced during the imaging and recording processes. Finally, the images were randomly rotated within an angle range from 0° to 360°. The authors noted that the rotation operation might cause some pixels to be rotated out of the image frame, leaving behind empty pixels within the frame. These empty pixels were filled with zero values. Finally, In the data pre-processing stage, Ouf [19] resized all the images in the dataset to 416 × 416 pixels to ensure compatibility with the Faster R-CNN and YOLOv4 models, avoiding memory constraints, low speed, and low accuracy issues. The image annotations, generated using the LabelImg tool, were saved in both .txt format (for YOLO) and .xml format (for PASCAL VOC dataset, which can be easily converted to TFRecords). For the YOLOv4 model, the images and annotations were directly input into the model using the Darknet deep learning framework. For the Faster R-CNN model, which utilizes the TensorFlow deep learning framework, the .xml annotations were converted into the TFRecord data type before being fed into the model.

The data preprocessing techniques discussed in this section underscore the significance of meticulous data preparation for achieving optimal performance in seed classification and germination studies. Researchers have employed a wide range of preprocessing methods, including image segmentation, contrast enhancement, noise removal, data augmentation, and image resizing, among others. These techniques aim to enhance image quality, extract relevant features, and ensure compatibility with the employed models. For future studies, a recommended preprocessing pipeline could involve a combination of techniques, such as nCDA or clustering methods like DBSCAN to isolate individual seeds from the background. This can be followed by image augmentation techniques (e.g., rotation, flipping, zooming) to increase the diversity of the training data and improve model generalization. Additionally, contrast enhancement methods like CLAHE can be applied to improve image quality, and noise removal techniques like median filtering can help mitigate the effects of noise introduced during image acquisition. Finally, image resizing to a consistent size compatible with the chosen model architecture can ensure efficient processing and prevent memory constraints.

7. Algorithms and Results

This section delves into presenting diverse methods, ranging from traditional machine learning techniques to sophisticated deep learning models, in these ten works. Here, I review the significant contributions and performance outcomes of these research studies, highlighting their effectiveness in improving seed classification accuracy and efficiency.

Table 3 Summary of Algorithms and Results from ten Studies on Seed Classification

SL.	Authors	Algorithm(s) Used	Own Architecture?	Results
1	Fu et al.	PCA, LDA, SVM	No	SVM: 92.71% accuracy
2	Gulzar et al.	Modified VGG16	Yes	Modified VGG16: 99% validation accuracy
3	He et al.	SVM, kNN, CNN, MobileNet, PointNet	No	Improved voting method: 97.4% average accuracy

4	Kiratiratanapr k et al.	LR, LDA, k-NN, SVM, VGG16, VGG19, Xception, InceptionV3, InceptionResNetV2	No	InceptionResNetV2: 95.15% highest accuracy
5	Loddo et al.	SeedNet, Comparison with other CNNs and traditional ML	Yes	SeedNet: 97.47% accuracy
6	Rathnayake et al.	Gradient-boosting algorithms (XGBoost, CatBoost, Light-GBM) with Cascaded-ANFIS	Yes	Cascaded-ANFIS: 76.97% accuracy for variety classification
7	Xu et al.	P-ResNet, Comparison with other CNNs	Yes	P-ResNet: 99.70% highest classification accuracy
8	Yasam et al.	MLRCM-SG (CLAHE, SIFT, RF, DT)	No	RF classifier: 92.65% accuracy
9	Zhao et al.	YOLO-r (includes image partitioning, Transformer encoder, small target detection layer, CDIoU loss)	Yes	YOLO-r: 95.39% mAP
10	Ouf et al.	Faster R-CNN with Inceptionv2, YOLOv4 with CSPDark-net53	No	YOLOv4: 98.52% mAP, 47.2 ms inference speed

Fu et al. [10] applied PCA, LDA, and SVM algorithms to the morphological, spectral, and combined datasets for classifying oat cultivars. PCA provided some grouping ability but could not clearly distinguish all 16 cultivars. LDA achieved 89.69% accuracy on the testing set using combined data, while SVM attained 92.71% accuracy. The results demonstrated the effectiveness of integrating multispectral imaging with multivariate analysis for rapid, non-destructive oat cultivar identification based on seed characteristics, providing a potential alternative to manual inspection and destructive molecular marker-based methods. Gulzar et al. [11] implemented a modified VGG16 architecture with five additional layers for seed classification. The model was trained using transfer learning and fine-tuned with an adjustable learning rate and model checkpointing. The training accuracy reached 99.6%, and the validation accuracy stabilized at around 99% after 40 iterations. The proposed model outperformed traditional machine learning algorithms such as KNN, decision trees, and random forests. The test results on unseen real-world data showed an accuracy of 99%, with minor misclassifications due to similarities in seed textures. The proposed model's performance was encouraging for industrial applications in seed sorting and packaging.

He et al. [12] used SVM, kNN, CNN, and MobileNet for 2D image classification, and PointNet for 3D point cloud classification. They proposed an improved voting method for late fusion, which generates a scoring vector based on each model's performance. The predicted probabilities from each model were weighted using the scoring vector, and the class with the highest probability was selected as the final prediction. The improved voting method achieved an average accuracy of 97.4%, outperforming individual models by 4.9% to 18.1%. The method significantly improved the identification accuracy of Hannuo35 and Yuanhan35 rice varieties compared to using only 2D models. Kiratiratanapruk et al. [13] applied four traditional machine learning methods (LR, LDA, k-NN, and SVM) and five deep learning models (VGG16, VGG19, Xception, InceptionV3, and InceptionResNetV2) for rice variety classification. The best accuracy achieved by the SVM method was 90.61%, 82.71%, and 83.9% in subgroups 1, 2, and the collective group, respectively. The deep learning techniques outperformed traditional methods, with InceptionResNetV2 achieving the highest accuracy at 95.15%. The authors also demonstrated that their seed orientation method improved classification accuracy by 1.3% in the deep learning experiment, and seed quality screening enhanced accuracy by 2-3% in statistical methods.

Loddo et al. [14] proposed SeedNet, a CNN architecture for seed classification and retrieval, and compared its performance with ten state-of-the-art CNNs and four traditional machine learning methods using handcrafted features. SeedNet outperformed all other methods in both classification and retrieval tasks on the Canadian and local datasets, achieving accuracies of 95.24% and 97.47%, respectively. The study also found that CNN descriptors generally outperformed handcrafted descriptors in both tasks. The authors concluded that SeedNet is robust, efficient, and accurate for seed classification and retrieval, and deep learning techniques have substantial potential in this domain. Rathnayake et al. [15] proposed a novel machine learning algorithm combining gradient-boosting algorithms (XGBoost, CatBoost, and LightGBM) with a Cascaded-ANFIS structure for classifying rice seed varieties and ages. The proposed algorithm outperformed 13 other feature-based machine learning algorithms in terms of accuracy, precision, recall, and F1-score for both variety and age classification tasks. The algorithm achieved an accuracy of 0.7697 for variety classification and accuracies ranging from 0.7551 to 0.9512 for age classification of individual varieties.

Xu et al. [16] proposed an improved convolutional neural network (CNN) architecture called P-ResNet and compared its performance with other popular CNN models, such as AlexNet, VGGNet, GoogLeNet, MobileNet, DenseNet, ShuffleNet, and EfficientNet. The authors employed transfer learning to train the models and evaluated their classification accuracy using the maize seed dataset. The P-ResNet model achieved the highest classification accuracy of 99.70%, outperforming other CNN models. The results demonstrate the effectiveness of the proposed P-ResNet architecture, transfer learning, and data augmentation techniques for the classification of maize seeds using machine vision and deep learning. Zhao et al. [18] proposed YOLO-r, an improved version of YOLOv5, for detecting the germination rate of rice seeds. They incorporated image partitioning, the Transformer encoder, a small target detection layer, and the CDIoU loss function to enhance the model's performance. The experimental results showed that YOLO-r achieved a mean average precision (mAP) of 0.9539, outperforming 12 other models. The average detection time per image was 0.011 seconds, meeting real-time requirements. The evaluation of YOLO-r on the RiceSeed2 dataset revealed that the mean absolute error of the germination rate was less than 0.1.

Ouf [19] employed Faster R-CNN with Inceptionv2 backbone and YOLOv4 with CSPDarknet53 backbone for leguminous seed detection. YOLOv4 utilized mosaic data augmentation to improve small seed detection. The models were evaluated using mean Average Precision (mAP) and inference speed. YOLOv4 outperformed Faster R-CNN with a mAP of 98.52% compared to 84.56%, and an inference speed of 47.2 ms versus 53.1 ms. YOLOv4 demonstrated superior accuracy and runtime, with an error rate below 2% and a running time under 2 seconds, making it an effective tool for leguminous seed detection in complex scenarios.

These studies demonstrate the significant potential of CNN and custom architectures for seed classification and identification. Particularly, works such as those by Gulzar et al. [11] and Loddo et al. [14], which implemented their own modified CNN architectures, have shown superior performance in terms of accuracy and efficiency compared to traditional machine learning methods and standard CNN models. These custom models, tailored specifically to the intricacies of seed imagery, capitalize on the unique features of seed morphology better than generic algorithms. Therefore, it is recommended for future research in seed classification to consider developing and employing bespoke CNN architectures. This approach not only enhances classification accuracy but also optimizes the models to handle the specific challenges posed by different seed types and imaging conditions, providing a robust solution for both academic research and practical applications in agriculture.

Limitations

This section outlines the limitations identified in these ten studies and recommends potential solutions to enhance model performance and applicability. Each cited work's specific challenges are discussed, with suggestions aimed at improving dataset diversity, model generalization, and evaluation metrics.

Table 4 Summary of Limitations and Recommended Solutions for ten Studies

Work	Limitation	Recommended Solution
Loddo et al.	Model limited to specific datasets with small number of classes and samples.	Expand dataset diversity and size; incorporate additional evaluation metrics like computational efficiency.
Gulzar et al.	Dataset limited to 14 seed types, potentially causing overfitting.	Increase dataset size and diversity; compare with state-of-the-art models.
Kiratiratanaprk et al.	Focus on Thai rice broader applicability. cultivars, limiting	Test model on diverse crop types and environmental conditions; explore data augmentation and transfer learning.
Fu et al.	Study limited to 16 oat cultivars, with no exploration of environmental factors.	Include diverse oat cultivars and study the impact of environmental factors on seed morphology.
Ouf et al.	Focus on leguminous seeds with a small dataset, limiting generalization.	Increase dataset size and diversity; explore environmental impacts on model performance.
He et al.	Limited scope to eight Chinese rice varieties.	Expand study to include diverse rice varieties from various regions.
Xu et al.	Limited to five maize varieties; does not consider seed age or storage conditions.	Include diverse maize varieties and study impact of seed age and storage conditions on classification.

Zhao et al.	Focused on rice seeds from Shanghai, which may not represent other regions.	Expand dataset to include diverse rice cultivars from different regions; address seed contamination challenges.
Rathnayake et al.	Limited to six Japanese rice varieties; does not consider environmental factors.	Expand dataset to include diverse rice varieties and study impact of environmental factors on age classification.
Yasam et al.	Limited to three crop types; does not address seed contamination or overlapping issues.	Expand study to include more crop types; enhance model to handle contamination and overlapping seeds.

The work by Loddo et al. [14] focuses on two specific datasets, which may limit the generalization of their proposed model to other seed types or datasets. Additionally, the number of classes and samples in their datasets is relatively small, potentially affecting the model's performance on larger and more diverse datasets. Furthermore, the evaluation metrics used primarily focus on classification accuracy, neglecting other important aspects such as computational efficiency and real-time performance. Gulzar et al.'s study [11] is limited to a dataset consisting of only 14 types of seeds, which may not be representative of the vast diversity of seeds encountered in practical applications. Additionally, the dataset's size, with around 2,000 images per class, is relatively small for deep learning models, potentially leading to overfitting or suboptimal performance. Furthermore, the authors do not provide a comprehensive comparison with other state-of-the-art models or alternative approaches.

The work by Kiratiratanapruk et al. [13] is focused solely on Thai rice cultivars, which may limit its applicability to other regions or crops. Additionally, the authors do not provide a detailed analysis of the model's performance on individual cultivars or specific scenarios, such as varying lighting conditions or seed densities. Furthermore, the study does not explore the potential impact of data augmentation techniques or transfer learning approaches, which could potentially improve the model's performance. Fu et al.'s study [10] is limited to 16 oat cultivars, which may not be representative of the diversity of oat cultivars globally. The authors do not explore the potential impact of environmental or processing factors on seed morphology and spectral characteristics, which could affect the model's performance in real-world scenarios. Additionally, the study does not provide a comprehensive comparison with other machine learning or deep learning approaches for seed classification and identification.

Ouf's work [19] focuses on object detection for leguminous seeds, which may not be directly applicable to other types of seeds or agricultural products. The dataset used is relatively small, with only 828 images containing 9,801 seed objects, potentially limiting the model's generalization capabilities. Furthermore, the study does not explore the potential impact of environmental factors, such as lighting conditions or seed maturity, on the model's performance, nor does it provide a detailed analysis of the model's computational efficiency or resource requirements. He et al. [12] primarily focuses on the classification of rice seed varieties using a multi-modal approach. However, the study is limited in its scope as it only considers eight rice varieties from China. The performance of the proposed method may vary when applied to a larger and more diverse set of rice varieties from different regions. Additionally, the study does not address the potential impact of environmental factors or storage conditions on the seed features, which could influence the classification accuracy.

Xu et al. [16] proposes a machine vision and deep learning approach for the classification of maize seeds. While the study demonstrates promising results, it is limited to five maize varieties commonly found in China. The generalization capabilities of the proposed method may be restricted when applied to a broader range of maize varieties from different geographical regions or with varying morphological characteristics. Furthermore, the study does not consider the potential influence of seed age or storage conditions on the classification performance. Zhao et al.'s work [18] focuses on developing a deep learning model (YOLO-r) for detecting rice seed germination and evaluating the germination rate. However, the study is limited to a specific dataset consisting of rice seeds from Shanghai, China. The performance of the proposed model may vary when applied to rice seeds from different regions or cultivars, as seed morphology and germination patterns can be influenced by environmental and genetic factors. Additionally, the study does not address the potential challenges posed by factors such as seed contamination or the presence of debris in the images.

Rathnayake et al. [15] proposes a machine learning model for the age classification of Japanese rice seeds. While the study introduces a novel dataset based on harvested age, it is limited to six rice varieties commonly found in Japan. The performance of the proposed model may be affected when applied to a broader range of rice varieties or seeds from different geographical regions. Furthermore, the study does not consider the potential impact of storage conditions or environmental factors on the seed features and age classification accuracy.

The limitations and recommended solutions for various seed studies are summarized in Table 4. This table provides an overview of the specific challenges faced in each cited study and proposes potential approaches to enhance the generalizability and effectiveness of seed classification and identification models.

8. Discussion and Recommendation

This comprehensive review has examined ten experimental research works on seed detection and classification using artificial intelligence techniques, particularly machine learning and deep learning algorithms. The survey has highlighted the significant potential of AI in advancing seed identification and categorization processes, which play a crucial role in the agriculture domain.

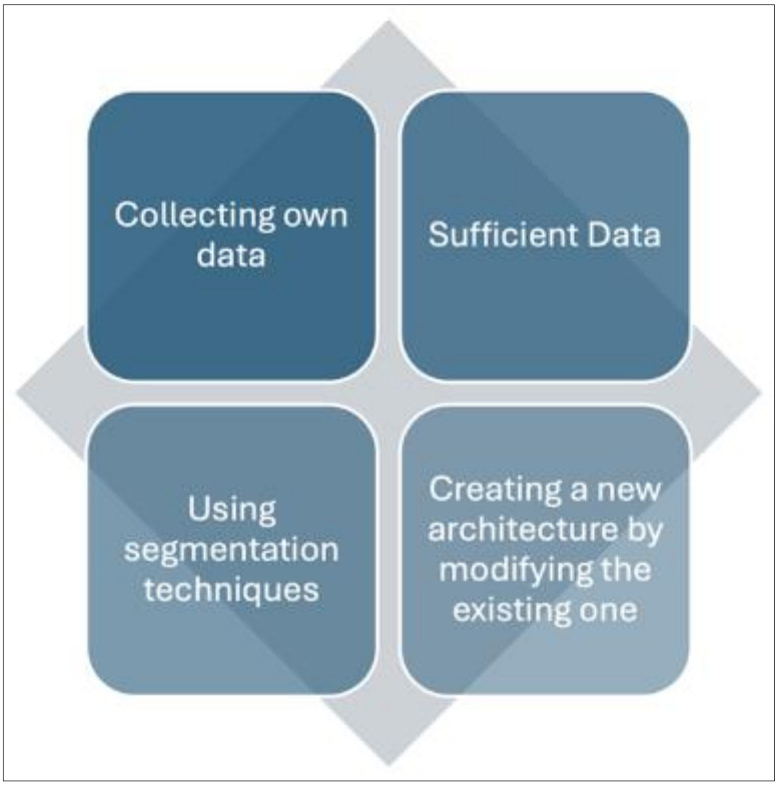

Figure 2 Recommendations for future work

Based on the in-depth analysis of these studies, the following four recommendations (see fig. 2) are proposed for future experimental analyses of seed detection and classification:

- Collecting Own Data: It is recommended that researchers consider collecting their own seed datasets tailored to their specific research objectives and target seed varieties. Many of the reviewed studies were limited by the use of existing datasets, which may not fully represent the diversity of seed types or environmental conditions encountered in real-world scenarios. By collecting their own data, researchers can ensure that their datasets capture the unique characteristics and variations of the seeds under investigation, leading to more accurate and generalizable models.
- Collecting Sufficient and Big Data: Several studies in this review were constrained by the limited size of their datasets, which can adversely impact the performance of deep learning models and their ability to generalize effectively. Future research should prioritize the collection of large and diverse datasets, encompassing a wide range of seed varieties, imaging modalities, and environmental conditions. Larger datasets can leverage the full potential of deep learning algorithms and enable the development of more robust and accurate models for seed classification and identification.
- Using Segmentation Techniques: Many of the reviewed studies employed image segmentation techniques to isolate individual seeds from the background or to separate overlapping seeds. Accurate segmentation is crucial for effective feature extraction and classification. Future research should explore advanced segmentation techniques, such as normalized canonical discriminant analysis (nCDA), clustering algorithms (e.g., DBSCAN),

or deep learning-based segmentation models, to improve the quality of input data and enhance the overall performance of seed classification models.

- Creating a New Architecture: Several studies in this review demonstrated the advantages of developing custom deep learning architectures tailored specifically for seed classification tasks. These architectures, such as SeedNet, P-ResNet, and YOLO-r, outperformed generic models by capitalizing on the unique characteristics of seed imagery. Future research should consider designing and implementing novel deep learning architectures that can effectively capture the intricate morphological and spectral features of seeds, leading to improved classification accuracy and efficiency.

By following these recommendations, researchers can address the limitations identified in the reviewed studies and contribute to the advancement of AI-driven seed detection and classification techniques. Collecting high-quality and diverse datasets, employing effective segmentation methods, and developing specialized deep learning architectures will enable the creation of more accurate, robust, and reliable models, ultimately benefiting the agriculture sector through improved seed quality control, inventory management, and crop productivity.

9. Conclusion

This study reviewed various AI-driven methods for seed detection and classification, highlighting significant advancements achieved through deep learning techniques. The survey revealed that Convolutional Neural Networks (CNNs) are predominantly employed due to their high accuracy and efficiency in processing seed images. Key limitations in existing studies were identified, including limited dataset diversity and generalization challenges, along with recommendations for future research. The research underscores AI's potential to enhance agricultural practices by enabling more accurate and efficient seed identification processes, ultimately contributing to improved crop quality and yield. Moving forward, insights from this study will guide the development of more robust AI models for widespread application in agriculture, benefiting society through sustainable food production advancements.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] V. J. Erickson and A. Halford, Seed planning, sourcing, and procurement, Restor. Ecol., vol. 28, no. S3, Aug. 2020, doi: 10.1111/rec.13199.

[2] T. Islam, T. T. Sarker, K. R. Ahmed, and N. Lakhssassi, Detection and classification of cannabis seeds using RetinaNet and Faster R-CNN, Seeds, vol. 3, no. 3, pp. 456–478, Aug. 2024.

[3] S. Y. Liu, Artificial Intelligence (AI) in Agriculture, IT Prof., vol. 22, no. 3, pp. 14–15, May 2020.

[4] W. Dana and W. Ivo, Computer image analysis of seed shape and seed color for flax cultivar description, Comput. Electron. Agric., vol. 61, no. 2, pp. 126–135, May 2008.

[5] T. T. Sarker, T. Islam, and K. R. Ahmed, Cannabis seed variant detection using Faster R-CNN, 2024, doi: 10.48550/ARXIV.2403.10722.

[6] L. Wang, M. Tanveer, H. Wang, and M. B. Arnao, Melatonin as a key regulator in seed germination under abiotic stress, J. Pineal Res., vol. 76, no. 1, p. e12937, Jan. 2024.

[7] L. K. Paudel, R. Acharya, S. Pokharel, S. Shrestha, and C. B. Bohara, Evaluating the effects of nitrogen application on the agricultural performance of improved rice varieties in Marin, Sindhuli, Nepal, World J. Adv. Res. Rev., vol. 23, no. 1, pp. 1893–1901, Jul. 2024.

[8] M. Wakchaure, B. K. Patle, and A. K. Mahindrakar, Application of AI techniques and robotics in agriculture: A review, Artif. Intell. Life Sci., vol. 3, no. 100057, p. 100057, Dec. 2023.

[9] T. Islam et al., Predictive modeling for breast cancer classification in the context of Bangladeshi patients by use of machine learning approach with explainable AI, Sci. Rep., vol. 14, no. 1, p. 8487, Apr. 2024.

[10] X. Fu, M. Bai, Y. Xu, T. Wang, Z. Hui, and X. Hu, Cultivars identification of oat (L.) seed multispectral imaging analysis, Front. Plant Sci., vol. 14, p. 1113535, Feb. 2023.

[11] Y. Gulzar, Y. Hamid, A. B. Soomro, A. A. Alwan, and L. Journaux, A convolution neural network-based seed classification system, Symmetry , vol. 12, no. 12, p. 2018, Dec. 2020.

[12] X. He et al., Multi-modal late fusion rice seed variety classification based on an improved voting method, Collect. FAO Agric., vol. 13, no. 3, p. 597, Mar. 2023.

[13] K. Kiratiratanapruk et al., Development of paddy rice seed classification process using machine learning techniques for automatic grading machine, J. Sens. Stud., vol. 2020, pp. 1–14, Jul. 2020.

[14] A. Loddo, M. Loddo, and C. Di Ruberto, A novel deep learning based approach for seed image classification and retrieval, Comput. Electron. Agric., vol. 187, no. 106269, p. 106269, Aug. 2021.

[15] N. Rathnayake, A. Miyazaki, T. L. Dang, and Y. Hoshino, Age Classification of Rice Seeds in Japan Using Gradient-Boosting and ANFIS Algorithms, Sensors , vol. 23, no. 5, Mar. 2023, doi: 10.3390/s23052828.

[16] P. Xu, Q. Tan, Y. Zhang, X. Zha, S. Yang, and R. Yang, Research on maize seed classification and recognition based on machine vision and deep learning, Collect. FAO Agric., vol. 12, no. 2, p. 232, Feb. 2022.

[17] Q. Peng, L. Tu, Y. Wu, Z. Yu, G. Tang, and W. Song, Automatic monitoring system for seed germination test based on deep learning, J. Electr. Comput. Eng., vol. 2022, pp. 1–15, Oct. 2022.

[18] J. Zhao et al., Deep-learning-based automatic evaluation of rice seed germination rate, J. Sci. Food Agric., vol. 103, no. 4, pp. 1912–1924, Mar. 2023.

[19] N. S. Ouf, Leguminous seeds detection based on convolutional neural networks: Comparison of Faster R-CNN and YOLOv4 on a small custom dataset, Artificial Intelligence in Agriculture, vol. 8, pp. 30–45, Jun. 2023.

[20] T. Godin et al., Real time noise and wavelength correlations in octave-spanning supercontinuum generation, Opt. Express, vol. 21, no. 15, pp. 18452–18460, Jul. 2013.

[21] M. T. Islam, T. Ahmed, A. B. M. Raihanur Rashid, T. Islam, M. S. Rahman, and M. Tarek Habib, Convolutional neural network based partial face detection, in 2022 IEEE 7th International conference for Convergence in Technology (I2CT), IEEE, Apr. 2022. doi: 10.1109/i2ct54291.2022.9825259.

[22] H. Duan, L. Mei, J. Wang, L. Song, and N. Liu, A new imaging model of Lytro light field camera and its calibration, Neurocomputing, vol. 328, pp. 189–194, Feb. 2019.

[23] T. Islam, A. Vuyia, M. Hasan, and M. M. Rana, Cardiovascular disease prediction using machine learning approaches, in 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), IEEE, Apr. 2023. doi: 10.1109/cises58720.2023.10183490.

[24] M. Dai et al., Green control for inhibiting growth by stress factors in forage grass factory, Front. Microbiol., vol. 15, p. 1437799, Aug. 2024.

[25] Z. Y. Xin, Z. H. He, Y. Z. Ma, and Q. S. Zhuang, Crop breeding and biotechnology in China, J. New Seeds, vol. 1, no. 1, pp. 67–80, Jan. 1999.

[26] M. Aljabri, M. AlAmir, M. AlGhamdi, M. Abdel-Mottaleb, and F. Collado-Mesa, Towards a better understanding of annotation tools for medical imaging: a survey, Multimed. Tools Appl., vol. 81, no. 18, pp. 25877–25911, Mar. 2022.