SOLENT UNIVERCITY, SOUTHAMPTON

FACULTY OF BUSINESS LAW AND DIGITAL TECHNOLOGY

MASTER IN APPLIED AI AND DATA SCIENCE 2022

“CROP SEEDS CLASSIFICATION”

STUDENT: BHUMILKUMAR PATEL

COURSE: DISSERTATION(COM726)

SUPERVISIOR: MUNTASIR Al-AFSOOR

Table of Contents

List of Figure

Abstract

Crop production is crucial in nations where agriculture accounts for a sizable component of the national economy. Especially among communities of farmers who spend the majority of their time working in fields and producing crops to restore their nation's economy, agricultural production serves as a backbone. Additionally, determining the kind of a crop's seed is essential to predicting the future output. The classification of a seed is a challenging task that farmers and agricultural workers sometimes find boring. Their job may occasionally be ineffective owing to errors in judgment or because of weather conditions, which undoubtedly halts both them and the national economy. Technology has dominated every industry in the world today, including transportation, manufacturing, agriculture, and finance. Systems that are supported by technology are crucial because they reduce the need for human labor and the likelihood of error thanks to their efficacy and efficiency. In light of this, the agricultural industry needs an intelligent system that automatically determines a seed's class based on a variety of characteristics. Modern technologies have been created and developed in the disciplines of artificial intelligence (AI) and deep learning (DL), which not only automate routine jobs but also allow computers to exercise intellect without human assistance. Through learning image features from seed images, this research will automatically assist farmers in the agricultural sector by addressing the classification of specific maize seeds. Kaggle, an online data science and machine learning platform with datasets from all industries, will provide the dataset for this study. Convolutional Neural Networks (CNNs) will be used to classify the seed pictures for the challenge. Future challenges for upcoming researchers will also be highlighted in this research.

Keywords: Artificial Intelligence (AI), Deep Learning (DL), Convolutional Neural Networks (CNNs), Kaggle, seed classification

Chapter 1

1. Introduction

1.1. Crop Production

The process of growing and caring for crops for domestic use might be characterised as crop production. The foundation of the economies of nations that rely largely on agricultural products is crop production. In South Asian countries like India, Pakistan, and Bangladesh, agriculture represents a substantial portion of the economy. To assist the farmer community in these countries, an automated system that can forecast crop yield in the agricultural sector should be created. The economy of these nations collapses if agricultural farming there declines due to factors like rainfall, malnutrition, a lack of

1.2. Seeds / Crop

A crop is often a plant that is grown or harvested for domestic use in order to make money and meet people's dietary needs (Merriam Webster Dictionary, 2017). A crop is defined as a big collection of identical plants growing in a selected region, such as a field. The bulk of crops are cultivated in agriculture. A crop may also include minute components like fungi. On a huge scale, crops including rice, sweet potatoes, bananas, ginger, black pepper, cashew nuts, and others are farmed. The rehabilitation and expansion of a county's economy, which is highly dependent on agriculture, depends on crop output.

1.3. Modern Technology Assessment in Agriculture Systems

In their research, K. Zhichkin et al. (2019) assert that in addition to agricultural production, a review of modern agricultural technology is also necessary. Evaluating the development of new technologies and equipment forms the basis for advancement in all sectors, including agriculture. Only under conditions of greater reproduction and on the foundation of scientific and technological advancements can industry preserve its long-term competitiveness. Government-to-government processes like upgrading farm machinery and the tractor park involve not only increases in productivity, cost savings in labour and logistical funds, process optimization, and immediate access to useful documentation, but also significant investment costs that may outweigh the potential rewards.

1.4. Research Contributions in Agriculture

Many automated technologies have been created to preserve and honor the hard work that farmers generate while keeping up with crop output and to help farmers and the agricultural industry. As can be seen from the explanation above, it emerges out that a sophisticated system must be created to help farmers estimate the crop yield in their farms. This field has been developed by several researchers, and machine learning (ML) and artificial intelligence (AI) have advanced to a level where they can effectively address it. Therefore, AI and machine learning (AI & ML) based systems have become necessities of the current day, notably in agriculture, which we'll touch on in the next paragraph, because they have decreased effort and time by adjusting to general intelligence and the interpretation of data.

1.5. Artificial Intelligence (AI)

Artificial Intelligence (AI) is the branch of computer science (CS) & is demanding and the most growing field in today's era. It is basically the science of making intelligent machines that could be able to portray human behaviour such as reasoning, understanding, listening, problem solving, etc. The researchers who work in this area of computer science are highly skilled and passionate. The applications of AI have sky as their limit, i.e., there is n sector in today's world in which artificial intelligence isn't being applied. From intelligent robots to automated production systems, speech recognition, image processing and self-driving cars, AI has conquered every sector with lots of notable applications. AI can also be applied to agriculture such as predicting crops production for future, intelligent classification of crop images, and so on. This research is aimed towards the intelligent prediction of crops in agricultural sector. AI is a big umbrella, underneath which are 3 key fields, i.e., machine learning & deep learning.

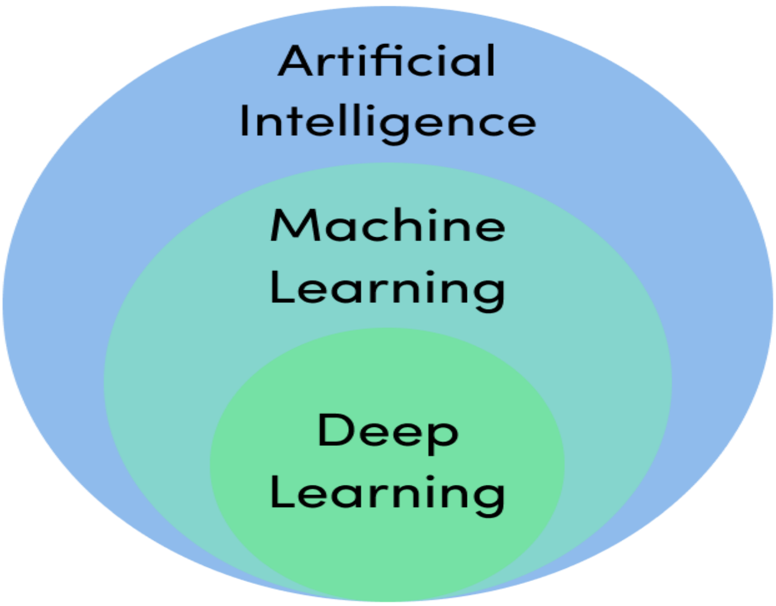

Figure 1 Hierarchy of Artificial Intelligence

1.6. Deep Neural Learning

Artificial neural networks, or broadly called as "deep learning," is a branch of machine learning that employs methods drawn from the functioning of the human brain (ANNs). The majority of ANNs used in deep learning operate in a layered framework, transmitting input from one layer to the next to extract more complex feature representations than traditional machine learning models. Convolutional neural networks (CNNs) are used to process images, recurrent neural networks (RNNs) are used to process text vectors, and so on. ANNs come in a variety of kinds, and each ANN is responsible for a certain set of tasks. Compared to traditional methods, ANNs possess a vast array of applications given the growing amount of data in the modern world. A picture shown below describes the variation between machine & deep learning.

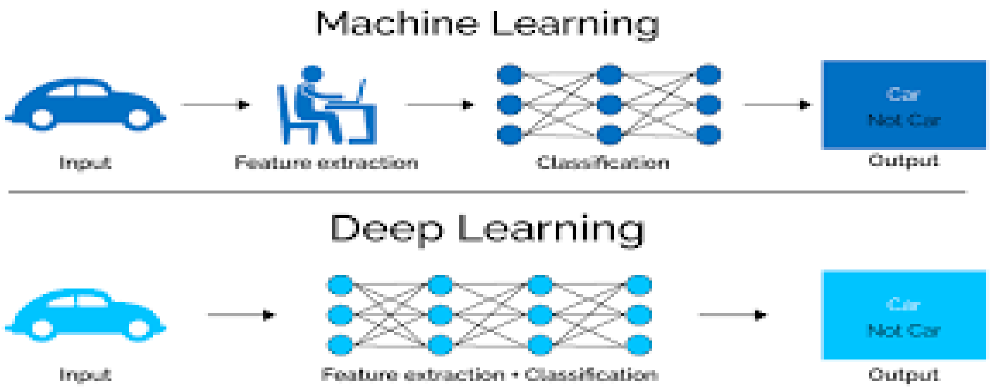

Figure 2: Comparison of Machine Learning & Deep Learning

1.7. Aim(s):

- To develop a deep learning model for intelligent classification of crops using deep learning.

1.8. Objectives:

- To assist farmers and agricultural staff in identification of seed of maize seed by making a deep learning CNN model
- To increase the accuracy of system to detect the class of a maize seed with minimum loss
- To figure out the class of a maize seed image based on convolutional neural networks (CNNs)

1.9. Research Questions:

The questions that motivate this research to proceed are as under:

- How to apply deep learning techniques in agricultural sector for intelligent classification of crop images?
- What is the significance of using deep learning models instead of classical models for classification?
- How making an intelligent model using artificial intelligence may aid agricultural sector to identify seed types?
- How will using a CNN instead of classical machine learning (ML) models will boost the performance of an image classifier?

1.10. Ethical Considerations

Ethics is a crucial subject to discuss, especially when doing efficient and error-free research. The UK Data Center Unit provides standards that include ethical issues for doing successful big data research, which also serves as the foundation for this study's ethics.

1.11. Project Plan

	ACTIVITIES	ASSIGNEE	EH	START	DUE	%
	Registration Form:		64h	16/May	27/May	100%
1	Analysis of Domain	Nishant Thakhr	16h	16/May	17/May	100%
2	Design of Problem Statement	Nishant Thakhr	24h	18/May	21/May	100%
3	Research Gap	Nishant Thakhr	24h	23/May	27/May	100%
	Research Proposal:		110h	09/Jun	27/Jun	100%
5	Problem Statement	Nishant Thakhr	8h	09/Jun	09/Jun	100%
6	Aim	Nishant Thakhr	8h	10/Jun	10/Jun	100%
7	Objectives	Nishant Thakhr	8h	11/Jun	11/Jun	100%
8	Research Question	Nishant Thakhr	8h	13/Jun	13/Jun	100%
9	Ethical Consideration	Nishant Thakhr	8h	14/Jun	14/Jun	100%
10	Literature Review	Nishant Thakhr	54h	15/Jun	23/Jun	100%
11	Methodology	Nishant Thakhr	8h	24/Jun	24/Jun	100%
12	Conclusion	Nishant Thakhr	8h	27/Jun	27/Jun	100%
	Literature Review:		168h	28/Jun	22/Jul	100%
14	Reading & Filtering Research Papers	Nishant Thakhr	96h	28/Jun	08/Jul	100%
15	comparison Research Paper	Nishant Thakhr	40h	11/Jul	15/Jul	100%
16	Purification of Research Writing Document	Nishant Thakhr	32h	18/Jul	22/Jul	100%
	Methodology:		80h	19/Jul	29/Jul	100%
18	Designing a Methodology	Nishant Thakhr	24h	19/Jul	21/Jul	100%
19	Selecting a Methodology	Nishant Thakhr	24h	22/Jul	25/Jul	100%
20	Finalization & Conclusion of Methodology	Nishant Thakhr	32h	26/Jul	29/Jul	100%
	Evaluation:		104h	04/Aug	19/Aug	100%
22	Data Finalization	Nishant Thakhr	16h	04/Aug	05/Aug	100%
23	Implementing Algorithm	Nishant Thakhr	32h	08/Aug	11/Aug	100%
24	Comparing & Evaluating Algorithm	Nishant Thakhr	24h	12/Aug	15/Aug	100%
25	Select Best Suited Model For Dataset	Nishant Thakhr	32h	16/Aug	19/Aug	100%
	Conclusion :		8h	22/Aug	22/Aug	100%
27	Conclusion	Nishant Thakhr	8h	22/Aug	22/Aug	100%
	Future Work:		8h	23/Aug	23/Aug	100%
29	Future Work	Nishant Thakhr	8h	23/Aug	23/Aug	100%
	Submission:		8h	25/Aug	25/Aug	100%
31	Submission		8h	25/Aug	25/Aug	100%

Figure 3: Project Plan

Chapter 2:

Literature Review

2.1 Classification of seed varieties using Deep Learning:

Sabanci et al. (2022) focused on pepper seed quality. The cultivar may impact growth procedures and circumstances. Pepper cultivars affect production and characteristics. Breeding projects may require proper seed cultivars. Pepper seeds may be differentiated visually. Small sizes and visual similarities make identifying seed varieties challenging. Computer vision and AI allow objective, quick cultivar discrimination. This study classified pepper seeds using CNN models. Green, orange, red, and yellow peppers provided the seeds. Pepper seed photos were scanned on a flatbed. After picture collecting, preprocessing, data augmentation, and deep learning-based categorization followed. Classification has two ways. ResNet18 and ResNet50 CNN models were trained on pepper seeds. In the second technique, pre-trained CNN model features were merged and then selected. SVM with different kernel functions classified all and selected characteristics (Linear, Quadratic, Cubic, and Gaussian). ResNet50 and ResNet18 have first-approximation accuracies of 98.05% and 97.07%, respectively. CNN-SVM-Cubic obtained 99.02% accuracy with specified features.

2.2 Machine Learning Based Agricultural Analysis:

Donmez (2022) developed machine-learning-based agricultural analysis systems. Computer-assisted methods are used in agriculture to boost output and product quality. Maize feeds the planet. Breeding time and output efficiency are affected by haploid and diploid maize seeds. Convolutional neural networks are used to classify haploid-diploid maize seeds (CNN). First, deep CNN layers were retrieved. Using MRMR (Max-Relevance and Min-Redundancy), CNN identified its top 100 features. These CNN features have been combined. Last stage of a typical classifier approach uses fused features for training and testing. According to experimental data, overall accuracy was 96.74%. The suggested method classifies maize seeds well.

2.3 Seeds Classification Using Machine Learning:

Koklu and Ozkan (2020) suggested a computer vision technique to classify seven registered dry bean types with comparable traits. Dry bean, the most widely grown culinary legume, has a great genetic variety. Seed quality affects agricultural yield. Seed categorization is vital for marketing and production to support sustainable agricultural systems. This study aims to give a strategy for producing consistent seed variants from population-based agricultural production, so seeds are not certified as a solitary variety. 13,611 grains of 7 registered dry beans were photographed for the model. MATLAB's GUI provides a user-friendly interface (GUI). CVS bean images were segmented and 16 features were retrieved (12 dimensions and 4 shape forms). 10-fold cross validation compared MLP, SVM, KNN, and DT classification models. MLP, SVM, KNN, and DT had 91.73, 93.13, 87.92, and 92.52% correct classification rates, respectively. The SVM classification model classified Barbunya, Bombay, Cali, Dermason, Horoz, Seker, and Sira beans with 92.36, 100, 103, 94.36, 94.92, 94.67%, and 86.84% accuracy, respectively. With these outcomes, farmers and buyers get consistent bean varietals.

2.4 Comparison of hybrid Seed Types Using Machine Learning Algorithms:

Nie et al. (2019) compared PLS-DA, SVM, and DCNN with hybrid seed types. Cross-breeding necessitates quick selection of acceptable hybrids. Selecting specific hybrid offspring is time-consuming and expensive. Near-infrared hyperspectral imaging and deep learning identified hybrid seeds. 6136 hybrid okra seeds and 4128 hybrid loofah seeds each had six variations. DCNN-based discriminant analysis has 95% accuracy. Final DCNN layer characteristics were exhibited using t-distribution stochastic neighbor embedding. The DCNN-based discriminant analysis approach reduced the labor and time necessary for cross-breeding progeny selection, accelerating related research.

Hyperspectral imaging is a non-destructive, quick approach for assessing seed quality and safety, according to Feng et al. (2019). This review offers uses for categorizing and grading, detecting viability and vigor, damage (defect and fungus), cleanness, and seed composition. Hyperspectral imaging technology is summarized for each category, including spectral range, sample variety, sample status, sample numbers, characteristics (spectrum features, picture features, and feature extraction techniques), signal mode, and data processing procedures. Hyperspectral imaging's use in seed quality and safety assessment indicates it can simplify ordinary operations.

Singh et al. (2020) said autonomic seed categorization can overcome manual problems. Agriculture needs it for practical and economic reasons. Researchers can create a more accurate data mining system using machine learning and AI. This paper proposes a new adaptive approach employing DIPS and FCRF. DIPS extracts area, perimeter, height, breadth, groove length, and asymmetry coefficient. Using these settings, the FCRF model classifies wheat seeds quickly. The technique helps classify seeds, separate damaged seeds, and manage seed quality based on grading standards. Experiments show that the suggested approach is more accurate than the existing wheat seed categorization algorithm. The proposed approach boosts performance by 97.7%.

Zhao et al. (2018) suggested discriminant models utilizing an RBFNN (RBFNN). Classifying seed varieties helps check purity and boost agricultural productivity. A hyperspectral imaging system was used to classify maize seeds. 12900 maize seeds from 3 kinds were tested. Extract and preprocess 975.01–1645.82 nm spectral data. Researchers explored how calibration sample size affects categorization accuracy. Results revealed that expanding calibration sample size marginally decreased calibration accuracy, but increased prediction accuracy. Optimal calibration set size was established. Principal component loading optimized wavelength selection (PCs). The RBFNN model designed with appropriate wavelengths and calibration set size achieved 93.85% calibration accuracy and 91.00% prediction accuracy. Applying this RBFNN model to each sample's average spectrum created a categorization map of seed varieties. Support vector machine verification was used to identify the best sample size (SVM). The results showed that hyperspectral imaging might be used to classify maize seeds and construct a real-time seed identification system.

2.5 Imaging and Spectroscopic Approaches for Seed:

Xia et al. (2019) described imaging and spectroscopic approaches for non-destructive agro-food quality assessment. Seeds are crucial to agriculture and forestry. Seed viability represents potential seed germination, and a speedy and effective approach to assess germination condition and seed viability before culture, sale, and planting is needed. Spectra and/or image processing and analysis are used to study seed quality. Seed quality is estimated using traditional and new image segmentation and spectral correction. This research examines near infrared, hyperspectral, multispectral, Raman, infrared thermography, and soft X-ray imaging for seed viability. Comparing basic theories, components, chemometric processing, analytical techniques, and prediction accuracies. The technological hurdles and future prognosis for these developing approaches are also highlighted.

2.6 Machine Vision Technology for seed Quality

Fazel-Niari et al. (2022) said machine vision technology is needed for cereal seed quality management. In this work, machine visual systems coupled with industrial digital cameras were used to identify and classify seven wheat grain classes. Study used two statistical models and three SVMs. 21,000 single grains were analyzed to establish their form, color, and texture. ReliefF ranked 91 characteristics. Shape, texture, and color were most prominent. The QSVM and first 35 features had the best classification accuracy. In a test utilizing independent data, this model's classification accuracy for sound white wheat, small white wheat, broken white wheat, shrunken white wheat, red wheat, barley, and rye was 98.7, 99.3, 90.7, 99, 100, and 97.3%, with an overall average of 97.6%. A machine vision system consisting of an industrial digital camera and quadratic support vector machine or non-linear discriminate analysis was utilised to explore the visual properties of wheat seeds.

2.7 Seeds Classification Using pixel-based CNN

Lv et al. (2019) suggested a pixel-based CNN to classify abstract deep features from VHRI. Pixel-based CNN requires much processing time and storage space. Super pixel CNN categorization has gained interest recently. It presents a CNN-based deep learning approach using SEEDS super pixels for VHRI classification. Three steps comprise the strategy. First, the image is segmented into homogeneous super pixels using GEOBIA to reduce processing units. Second, CNN's training and testing data are super pixelized. Third, training data is used to train CNN parameters and VHRI abstract deep features are retrieved. It classify two VHRI datasets using deep features. SEEDS and three additional super pixel segmentation methods were investigated to verify SEEDS-based CNN classification. CNN classification was best with super pixel extraction utilizing SEEDS. Comparing four super pixel segmentation algorithms exhibited CNN classification accuracy. For a simple VHRI with clear artificial objects and simple texture, several scales perform better than a single scale; for a complex VHRI with many complex objects, a single scale performs better.

Qiu et al. (2018) used hyperspectral imaging with CNN to detect rice seed variations. Four rice seed variants were imaged at 380–1030 nm and 874–1734 nm. Spectral data from 441–948 nm and 975–1646 nm were retrieved. KNN, SVM, and CNN models have various training sample sizes (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500 and 3000). Spectral range 2 KNN, SVM, and CNN models fared marginally better. Increased training samples enhanced model performance. Large training samples didn't increase performance. CNN model outperformed KNN and SVM in most situations, indicating its usefulness in analyzing spectral data. This study revealed that CNN may be used to analyze spectral data. Future studies must study more rice types to expand CNNs' application in spectral data interpretation.

2.8 Detectection of seeds using machine vision:

Bagheri and Heydari (2021) detected weed seeds using machine vision, ANN, and CDA. Amaranthus retroflexus, Amaranthus blitoides, Chenopodium album, Alyssum hirsutum, and Sinapis arvensis seeds were photographed. Images were processed to retrieve seed form characteristics. Raw and standardized data were extracted. Stepwise regression detected seed shape properties. The raw and standard ANN accuracy was 84.30% and 83.39%, respectively. ANN's raw and standard stepwise regression data identification accuracy was 84.30% and 83.39%. CDA technique findings indicated raw and standard data identification accuracy of 84.9% and 84.7%. This approach identified raw and standard stepwise regression data with 82.6% accuracy. A. retroflex and hirsutism had the best identification accuracy (>95%). In both approaches, C. album seed identification was over 87%. This shows that form characteristics in pattern recognition algorithms might identify weed seeds.

Shalabi et al. (2022) investigated 19 Ericaceous specimens from 10 genera. The study included seed macro, micromorphology and SDS-PAGE analyses. UPGMA-clustering approach employing NTSYS-pc 2.2 software expressed phonetic associations of examined taxa. The 47-character UPGMA phonogram separated A and B. (B). Group (A) separated into two sub ordinary clusters (AC) that expressed Vaccinoideae and (AD) that expressed Ericoideae with main group (B). These two subfamilies contain all studied genera. Vaccinieae, Gaultherieae, Oxydendreae, Lyoniae, and Andromedeae are Vaccinoideae tribes. Ericoideae has two clades, Phyllodoceae and Rhodoreae. Hierarchical taxonomy usually matches family classifications. Clustering of Menziesia pilosa with Rhododendron menziesii in close proximity with all Rhododendron species validated the placement of both genera under tribe Rhodoreae and supports the transfer of genus Menziesia to be nested in Rhododendron as proposed by recent DNA cladistics analysis.

Maize is one of the most widely grown grains, according to Altuntaş et al. (2019). Doubled-haploid shortens breeding time and enhances efficiency in advanced maize breeding, crop improvement, and genetic programmes. This breeding method requires choosing haploid seeds.

Manual labour consumes time. It employed computer vision to create a rapid, nondestructive model. CNNs automatically differentiated haploid from diploid maize seeds. Using AlexNet, VVGNet, GoogleNet, and ResNet. 1230 haploid and 1770 diploid maize seeds were utilised. R1-nj anthocyanin categorizes samples. It measured CNN models' accuracy, sensitivity, specificity, quality index, and F-score. VGG-19 was the best CNN model. VGG-19 has 94.22% accuracy, 94.58% sensitivity, 93.97% specificity, 94.27% quality index, and 93.07% F-score. Experiments revealed CNN models can recognize haploid maize seeds. This strategy is also superior to machine learning and manual selection.

Uzal et al. (2018) said commercial and research plant breeding operations require large-scale phenotyping. Manual phenotyping (costly, time-consuming and sometimes arbitrary). Computer vision can overcome some of these challenges. Deep Learning and CNNs have improved over time. In this work, it use computer vision to estimate soybean seed quantity, a process normally done by humans. To accomplish so, it used FE, SVM, and CNNs. It describe how to easily build up traditional CNNs and display the model's key features. It examined several season batches with both techniques and attained 50.4% (FE + SVM) and 86.2% (CNN) accuracy, suggesting the great generalization potential of a deep learning methodology in this study. Public data, code.

Tang et al. (2018) explored categorization and weakly guided localization of thoracic diseases using chest radiographs. A CNN-based AGCL system incorporates radiological severity levels. CNN receives tough visuals to boost learning. Highly confident samples (measured by classification probabilities) and their CNN-produced class-conditional heat maps are entered into the AGCL framework to learn additional unique convolutional features in the next iteration. In addition to classification, a two-path network architecture regresses heat maps from seed samples. Joint learning improves classification and localization and provides more seed samples for the following iteration. It test our iterative refining technique on the public ChestXray14 dataset. AGCL improves classification AUC by 5.7% (averaged across 14 illnesses) and localization recall/precision by 7%/ and 11%.

2.9 Deep convolutional neural networks use in Seed Classification

Deep convolutional neural networks were used by Singh et al. (2022) to create the model based on transfer learning for object detection and classification. Despite having been trained on large general-purpose datasets, application-specific datasets must still be manually labelled. When analyzing the purity of canola seeds, it takes into account this larger issue since end users want to tell out interesting species from contaminants in optical microscope pictures. A detector network trained to recognize seeds labels the dataset needed to train an analyzer network capable of seed identification and classification. To support our progressive strategy, it includes three experiments with 25 contaminant species, including Primary and Secondary Noxious Weed Seeds (as per the Canadian Weed Seeds Order). Literature assessment compares proposed system to rivals.

2.10 Machine Learning Based Approach

A machine learning-based approach was presented by Jiaoyang and Xingshi (2019) because machine learning is essential to data analysis. Support Vector Machine (SVM) is used to categorize and predict wheat seed data sets to identify them. This study uses experimental data from UCI Machine Learning Repository. Second, SVM with Lasso and elastic net regularization is used to categorize and forecast wheat seed. Experiments show SVM-Elastic Net's good classification accuracy.

Accelerated ageing is utilized to assess the physiological quality of soybean seeds, according to Matera et al. (2019). Few studies compare it to other crop physiological testing. This study examined the relationship between physiological performance measured by common vigour tests and accelerated ageing (germination, first count, seedling vigour classification, accelerated aging, electrical conductivity, emergence speed index, final emergence in sand substrate, and field emergence). Using different vigour seeds, a randomized experiment with four replications was conducted. Simple linear regression and Pearson simple correlation analysis (r) were conducted on the data. The lowest relationships with accelerated ageing were found for germination ($r = 0.8690$), initial count of germination ($r = 0.8460$), and electrical conductivity ($r = 0.8912$). The tests for seedling vigor classification and field emergence had the highest correlations with accelerated ageing ($r = 0.9160$ and 0.9198, respectively).

Seed quality is critical for consistent seedling establishment and excellent crop output, according to Thakur et al. (2022). This work suggests utilizing deep transfer learning (TL) with laser backscattering to automatically discover and classify high-quality seeds. The proposed sensor uses DL-based algorithms to analyze a single backscattered picture of a seed sample. The proposed sensor uses less memory, is resistant to external noise and vibration, requires less alignment effort, and has simple acquisition and processing units. Using DL-based processing frameworks such as CNN and other TL models (VGG16, VGG19, InceptionV3, and ResNet50) allows automated extraction of abstract features from pictures without additional image processing, improving classification effectiveness. All DL models performed well and were accurate, however InceptionV3 surpassed the others with 98.31% accuracy. Standard quality parameters including % imbibition, radicle length, and germination percentage (GP) were used to establish the sensor's performance. These metrics changed significantly (p 0.05), demonstrating the recommended sensor can better monitor seed quality. Additionally, the sensor is appropriate for real-time applications due to its experimental simplicity and DL-based automated categorization.

2.11 Spectrum Imaging Techniques

ElMasry et al. (2019) say spectrum imaging techniques have emerged to address quality evaluation conundrums in food and agriculture. This shows spectroscopy and imaging are complementary. Modern multispectral imaging and multivariate chemo metric analysis have been effectively used to food quality and safety management and seed science and technology. This is because of the ability to capture spatio-spectral data across the electromagnetic spectrum. This study provides an overview of all existing methods for acquiring, processing, and reproducing multispectral pictures for seed quality evaluation and seed phenotyping challenges. It also describes the basic design of the systems.

To satisfy the identification needs of accurate detection, Fu et al. (2022) examine a stacked sparse auto encoder paired with a cuckoo search (CS) optimized support vector machine (SSAE-CS-SVM). Identification of the many maize seed varieties is essential for raising maize output and quality. First, Savitzky-Golay (SG) together with conventional normal variables were used to analyze near-infrared (NIR) (871.61-1766.32 nm) hyperspectral data of maize seeds (SNV). Then, competitive adaptive reweighted sampling, SSAE, SAE, PCA, and principal component analysis (PCA) were used for feature extraction. Finally, the Soft Max regression model, SVM, and CS optimized SVM were used to build the recognition model (CS-SVM). The training set and testing set accuracies of the SSAE-CS-SVM model were 99.45% and 95.81%, respectively, according to the findings, indicating that it performed satisfactorily. This work demonstrated the enormous potential of merging deep learning algorithms and NIR hyperspectral imaging technologies for the identification of maize seed varieties.

According to Ramrez and Briceo (2022) Venezuelan plant groups' morphological fruit types were evaluated, as well as their richness and variety (measured by the Shannon-Wiener index) and fruit and seed colors. Almost all ecological measures analyzed indicated a positive association, including plant species. Richness in fruit types, fruit and seed colors, and Shannon-Weiner diversity score were all connected to plant species. Classification of plant locales indicated approximatively associations with geographic areas, vegetation types, and plant groups. This implies a predictable value and emphasizes how crucial fruit type composition, richness, and variety are to how plant communities are structured and how biotic and abiotic factors interact with one another.

2.12 Quantification of Seed Form in Various Species

According to Cervantes and Gomez (2018), the quantification of seed form in various species of the families that make up the order of Cucurbit ales is carried out by comparing photographs of the seeds with geometric shapes. A helpful technique in plant description for phenotypic characterization and taxonomic research is the quantification of seed form. The J index measures how much a seed's appearance resembles a geometric shape, and it is helpful in taxonomy when examining connections between different plant species. Three geometric shapes—the ovoid, two ellipses with various x/y ratios, and the contour of the Fibonacci spiral—are utilized as models in the Cucurbit aceae. The values of the J index were calculated after being compared to photographs of seeds. The findings for 29 species of the Cucurbitaceous family provide evidence for a connection between species ecology and seed morphology. Simple seed shapes, such as those that resemble basic geometric shapes like the ovoid, ellipse, or Fibonacci spiral, may be a characteristic of taxonomic groupings' basal clades.

2.13 Significant Role in Global Human Diet

Guerin et al. (2020) According to descriptions, palms have a significant role in global human diet, particularly as important suppliers of vegetable oils. However, it know little about the variety of lipids found in seeds and fruits belonging to the Arecaceae family. Therefore, our goal was to investigate the connections between the lipid content of seeds and fruits, the fatty acid composition of the corresponding tissues, phylogenetic variables, and biogeographical data. In the seeds and fruits of 174 and 144 different palm species, respectively, oil content and fatty acid composition were characterized. The chemotaxonomic importance of these features and their possible connection to ecological parameters might be assessed using distribution, linear regression, and multivariate analysis. It was discovered that there is a large intra-family variability for lipid characteristics. The tribe Cocoseae had the most lipid-rich species, whereas species that accumulated oil in the mesocarp were found in two-thirds of the tribes and all of the subfamilies. The lipid content of the fruit and seeds did not correspond. Mesocarp oil's fatty acid makeup varied greatly among tribes. In contrast, minimal within-tribe diversity and considerable between-tribe variability were observed for seed lipid characteristics. As a result, species belonging to the same tribe were grouped using multivariate analysis of seed lipid characteristics. Most palm species have a predominance of medium-chain fatty acids in their seeds, but in rare instances, these acids were also collected in the mesocarp. At the lowest natural latitude, seed unsaturated fatty acid concentration was associated with temperature. Numerous palms that had not previously been defined were discovered to be possible new sources of vegetable oils for use in food or other products. Seed lipid characteristics are very pertinent to palm chemotaxonomy because they indicate genetic drift that took place throughout the family's radiation. Our findings also point to the possibility that seed unsaturated fatty acids, which keep storage lipids liquid for effective mobilisation during germination, may offer an adaptation benefit in the coldest settings where palms are found.

2.14 Four Polar Solvents

Hasan (2022) used four polar solvents (petroleum ether, acetone, ethanol 70%, and distilled water) to extract parsley seeds and separate their phenolic components. HPLC analysis shows the plant contains phenolics (Campferol, Gallic acid, Quercetine, Rutin and Epigenine). Plant extracts and phenolic components reduce DPPH free radicals, according to study. The ethanolic extract at 500 and 400 g concentrations had a 96.2% and 94.8% difference from the control sample (ascorbic acid) at the same dosages. The study also showed that ethanoic extract was superior to other plant extracts and phenolic compounds. It also showed that the quantity of antioxidant chemicals increased the stability and inhibition of free radicals, which in turn prevented damage.

2.15 Seeds For Bioactive Components

Yuan-Yuan et al. (2022) tested SRA seeds for bioactive components. Ultra-performance liquid chromatography-electrospray ionisation identified 235 components in SRA seed. According to Wakholi et al. (2018), seed businesses and farmers both benefit from knowing the viability state of seeds before to planting (for yield prediction) (for seed warrant determination). However, a number of variables work together to diminish or entirely eliminate the viability of seeds both before and after harvest. Although several techniques have been used to assess the viability of seeds, hyperspectral imaging is possibly one of the most promising. This is due to its fast speed and capacity to identify the inside state of seeds without causing any damage, making it the ideal choice in particular for industrial sorting applications. Hyperspectral imaging will be used in this project to identify the most appropriate classification model or models for categorising maize seeds according to their viability. For this investigation, 600 samples of maize were chosen, with half receiving microwave heat treatment and the other half serving as the control group. All samples were scanned with a shortwave infrared hyperspectral camera having a wavelength range of 1000–2500 nm. LDA, PLS-DA, and SVMs were examined to determine the best classification model (SVM). The SVM model outperformed the prior study PLS-based technique by 5% and produced the maximum spectrum classification of up to 100%. The model also generated perfect classification pictures, indicating that viability-based categorization of maize using hyperspectral imaging may be possible. In conclusion, the findings of this study represent a significant advancement toward the creation of an efficient and non-destructive large-scale sorting system based on hyperspectral technology for the assessment of maize viability.

2.16 Perfection In Production Seed Quality

According to Pereira et al. (2020) seed businesses are striving for perfection in production quality by employing stringent procedures like the tetrazolium test (TZ test) and vigour definition. These procedures, however, are very time-consuming since they call for the expertise of a professional and the visual examination of a sizable sample of seeds in order to assess the strength of the seed lot. The TZ test has a set procedure, however because this analysis is a subjective human process, it may differ from analyst to analyst despite this. Numerous initiatives have been made in this regard to automate the analytical process in an effort to lessen inherent issues. As a result, this article offers strategies for understanding and categorising soybean seed vigour. To enhance the selection of the most instructive examples for the learning process, various active learning methodologies are also put forward. A thorough experimental assessment is conducted taking into account various datasets and cutting-edge learning methodologies. Based on the results, it is plausible to conclude that active learning methods produce classifiers that are more reliable and attain greater accuracies more quickly (in fewer learning iterations) than do classic supervised learning methods. Additionally, the number of labelled samples utilised in the learning process was reduced by 95.22%.

In the Malvaceae and other families of the order Malvales by Martn Gómez et al. (2019) looked at the form of the seeds. Comparing the seed shape to the cardioid helped to quantify it. When the J index is over 90, similarity is taken into consideration. The J index measures the degree of similarity between the two pictures, the seed and the cardioid. 73 genera of seeds were examined for seed form, and seeds with a cardioid-like shape were discovered in 10, with two of those genera belonging to the Bixaceae and two to the Cistaceae and two to the Malvaceae. Comparing 105 species' seeds to the cardioid allowed researchers to quantify seed morphology. The J index values were shown to be correlated with plant forms, with higher J index values.

2.17 Seed Variety Classification Using Deep Learning

Wu et al. (2021) Rapid seed variety classification helps breeders check for specific features and market regulators detect seed purity. High-quality, large-scale samples can be costly, making it challenging to create an appropriate classification model. This work used hyperspectral imaging (HSI) and deep transfer learning to classify crop seed kinds under sample-limited conditions. A sample-rich Pea dataset was used to design three deep neural networks. VGG-MODEL classified rice, oat, wheat, and cotton with limited samples with a 99.57% accuracy rate. Deep transferred model accuracy was 95, 99, 80.8, and 83.86% across four datasets. Deep transferred model always outperformed classic approaches with different-sized training sets. Visualization of deep characteristics and classification results showed the mobility of shared seed spectral properties, giving a way for rapid and reliable crop seed variety classification. HSI coupled with deep transfer learning outperformed other methods for sample-limited seed discovery. This study provides a fresh notion for agricultural germplasm screening under sample-limited conditions based on HSI.

2.18 DCNN-Based Discriminant Analysis

Nie et al. (2019) Cross-breeding requires speedy and effective selection of eligible hybrid offspring. Selecting specific hybrid offspring is time-consuming and expensive. Near-infrared hyperspectral imaging and deep learning identified hybrid seeds. 6136 hybrid okra seeds and 4128 hybrid loofah seeds each had six variations. PLS-DA, SVM, and DCNN were tested on hybrid seed types. DCNN-based discriminant analysis had 95% accuracy. Final DCNN layer features were exhibited using t-distribution stochastic neighbour embedding. The DCNN-based discriminant analysis approach reduced the labour and time necessary for cross-breeding progeny selection, accelerating related research.

Zhu et al. (2019) Cotton seed quality affects yield. Seven cotton seed variants were identified using near-infrared hyperspectral imaging. PCA score pictures demonstrated disparities across cotton seed kinds. PCA loadings determined effective wavelengths. Classification models were created using a CNN and ResNet. PLS-DA, LR, and SVM were used to classify complete spectra and effective wavelengths. Deep CNN and ResNet features were employed to classify cotton seeds using PLS-DA, LR, and SVM models. Deep features performed somewhat better than whole spectra and effective wavelengths in LR and PLS-DA models. Self-designed CNN models outperformed ResNet models. Most classification models employing complete spectra have calibration, validation, and prediction accuracy > 80%. Near-infrared hyperspectral imaging using deep learning can detect cotton seed variants.

Torres-Cortés et al. (2018) sequenced the shotgun DNA of microbial assemblages present on bean and radish seeds, germination seeds, and seedlings. After germination, there was an increase in available nutrients may have caused this selection. In contrast to bacteria isolated from seeds, taxa picked in seedlings showed a faster rate of bacterial growth. Our findings reveal that while there are variations in seed microbiota between plant species, food availability during germination induces changes in the structure of microbial communities. The evidence given here increases our understanding of the plant microbiome by evaluating changes in the microbial community during plant emergence.

2.19 Disease Detection in Agriculture

According to Coysh and Mead (2022) studies neurodegenerative illnesses are caused by prion-like host protein misfolding. RT-QuIC and PMCA assays have revolutionized the investigation and detection of prion disease and other neurodegenerative illnesses. Neurodegenerative disease clinical trials are expensive, slow, and ineffective. Seed amplification techniques in clinical trials can identify in vivo proteopathic seeds crucial to neurodegenerative disorders. Crop production, which is a key factor in every nation's economy, was the focus by Kurmi and Gangwar (2022). Plant disease detection is a major aspect in agriculture that must be kept under control to preserve a nation's economic progress. The suggested algorithm's contribution is to improve the extracted data from the existing resources such that the outcome is better without adding any unnecessary complexity. Prior to classifying an image as healthy or unhealthy, the suggested approach essentially localizes the leaf region. This approach is new in that it combines information that has been gathered from the available resources and optimizes it to improve the anticipated result. In order to identify the seed area, the leaf colours are investigated using color transformation. Mapping a low-dimensional RGB image to L*a*b increases the spectral range. Early seeds are affected by surrounding leaf development. The recommended technique has a 0.932 AUC and 0.903 accuracy.

2.20 Traditional Phenotypic Analysis Techniques

Amezquita et al. (2022) Traditional phenotypic analysis techniques are described and certain traits are measured, but the information encoded in shape is not fully measured. Biology is fundamentally shaped by shape. Using topological data analysis (TDA), specifically the Euler characteristic transform, it may extract, compare, and analyze this underlying information. TDA measures shape using algebraic topological representations. 3121 barley seedlings were scanned using X-ray CT at 127 m resolution to assess its application. Conventional and topological shape descriptors are computed. The Euler characteristic transform analyses an object's topological features along axes to determine its shape. According to a Kruskal-Wallis study of the topological signature, the Euler characteristic transform captures the seeds' crease and bottom. Topological shape descriptors can cluster seeds by panicle, but conventional shape descriptors can only do so by accession. Based on grain shape, it trained an SVM to classify 28 barley accessions. It find that the classification of barley seeds is improved when traditional and topological descriptors are combined rather than when traditional descriptors are used only. This development shows that TDA is an effective addition to conventional morphometrics for fully describing a variety of "hidden" form details that are otherwise undetectable.

Le et al. (2021) looked at how air NBs affected the rate of muskmelon seed germination. In industries including agriculture, aquaculture, food processing, and wastewater treatment, nano-bubbles (NBs) show their full potential. Untreated distilled water served as the control. Distilled water was treated with NBs for 10, 15, and 20 minutes. Between treatments, there were variations in seed germination rate that were statistically significant. After 14 days of treatment with water containing NBs, growth indices such as stem height, root length, and leaf area were assessed. Although these variations were fairly minor, statistical analysis revealed that there were significant differences in the dry weight of samples across treatments.

Review by Khazaei et al. (2019) globally, grain legumes are largely acknowledged as important dietary protein sources. For food processing formulations, lentil seeds are a great source of plant-based proteins and a sensible substitute for soy and animal proteins. In addition to providing dietary amino acids, lentil proteins are a source of bioactive peptides that have positive health effects. This study focuses on the state-of-the-art understanding of lentil protein, bioactive peptides, extraction and isolation techniques, and culinary applications. The fastest-growing crop for direct human consumption is lentil, which also has higher potential as a source of protein for food processing. The nutritional value of this rapidly developing crop will be improved internationally as processing fractions, amino acid composition, and protein quality are improved in lentils.

2.21 Major Problem while Growing Seed

Polder et al. (2019) viral infections are a major problem while growing seed potatoes. Virus-infected plants are declassified or even rejected as seed lots once they are discovered in the field, which results in a financial loss. Farmers spend a lot of work into finding sick plants and removing them from the field. However, depending on the cultivar, plants with viral diseases may go undetected during eye examinations, particularly in the early stages of culture. Therefore, rapid and accurate illness diagnosis is required. Modern vision techniques can drastically cut expenses by detecting unhealthy plants early. Previous lab tests shown that hyperspectral imaging could clearly discriminate between virus-infected and uninfected potato plants. This paper details our first substantial field study. A novel imaging system with a hyperspectral line-scan camera was created. A 5 mm line interval was used for the hyperspectral pictures that were captured in the field. For hyperspectral pictures, a fully convolutional neural network was modified and trained on two field experiments. Two more rows with various potato cultivars were used to validate the trained network. The accuracy and recall relative to traditional illness evaluation were more than 0.78 and 0.88 for three of the four row/date combinations. This demonstrates the method's usefulness for identifying diseases in the actual world.

2.22 Local Seed Supply Networks:

Local seed supply networks are important in popularizing wheat varietal innovations, according to Khed et al. (2021). Despite new technology being most economically advantageous for farmers, they do not adopt them. Most micro level studies make an attempt to address these adoption-related concerns, separating farm households from their larger environment, for instance, by ignoring the varied influences of local infrastructure, institutions, and legislation. Determine the importance of regional seed distribution networks in this study. According to conventional thinking, farmers' acceptance of new varieties and rates of varietal turnover are influenced by local seed supply networks. However, raises the possibility that the link may be more complicated and indirect. Farmers' preferences for current or vintage crop kinds may influence the seed supplier they choose. This link may have significant policy implications. With the current datasets, one can't definitively disprove the theory, thus it advise gathering more complex data as part of the AGG Wheat project's farm-household survey, including both qualitative and quantitative data. Supply chain, adoption, adoption of seed technology, and agriculture.

2.23 CNN Approches For Seed

Picon et al. (2019) described CNN's applicability in agronomy, especially for assessing plant visual symptoms. There is a choice between (1) creating smaller models for specific crops or (2) creating a unique multi-crop model, a much more difficult task (especially at early disease stages) with the benefit of the entire multi-crop image dataset variability to enrich image feature description learning as these models increase in both training images and supported crops and diseases. In this study, it give a large dataset of smartphone images taken outside in real-world scenarios. This collection shows wheat, barley, maize, rice, and rapeseed with several diseases. It achieved balanced accuracy (BAC=0.92) while building crop-specific models and balanced accuracy (BAC=0.93) when generating a single multi-crop model. This paper describes three CNN architectures that mix crop and non-image meta-data. This simplifies sickness classification and allows simultaneous multi-crop learning. The crop-conditional plant disease classification network outperforms all earlier systems and eliminates 71% of their miss-classifications. Concatenating context at the embedding vector level does this.

2.24 Deep Learning-Based Analysis in the Agriculture Sector

The preparation of a substantial amount of training data, as described by Toda et al. (2020), necessitates a time-consuming human data annotation procedure that frequently proves to be the limiting step. In this article, it demonstrate how a neural network for instance segmentation that aims to characterize the morphology of different cultivars of barley seeds may be adequately trained just using synthetic data. In comparison to the test dataset from the real world, the trained model had an average precision of 95% and a recall of 96%. It demonstrate the versatility of our method by demonstrating its efficacy with other crops, such as wheat, oats, lettuce, and rice. When using deep learning-based analysis in the agriculture sector, creating and using such synthetic data can be a potent way to reduce the expenses associated with human labor.

2.25 Classification On Weed Seeds Based on Images

(Luo et al., 2021) performed classification on weed seeds based on images, and applied deep learning for performing an intelligent classification. This research claims to be productive in developing a crop weed seed detection system, which would be able to be applied in various applications. In this research, the images of weeds in dataset are segmented and six convolutional neural network (CNN) models are used and compared with each other for identification of the best model possible for classifying species of weed seeds. Due to applying expansion on weed dataset, the applied CNN model can easily be retrained and the existing CNN model can be easily replaced for performing an upgrade to the existing model. For performing classification on weed seed images dataset, pre-trained CNNs are used such as AlexNet & GoogleNet, which are trained on training dataset with random selection and testing on 14,000+ samples, and the model with highest accuracy is selected, which is claimed to be GoogleNet in this research. In addition, this research also claims a more effective management of seed control by application of proposed seed detection model.

2.26 Deep Learning Based Approach

(de Medeiros et al., 2021) used a deep learning based approach on X-ray images dataset for performing classification on Crambe Abyssinica seed quality. This research claims that owing to the advancements of artificial intelligence (AI) in agriculture, significant improvements have been seen in this department. Advanced tools and technologies have been developed to identify the class of a seed to improve decision making in agriculture department. Keeping in context the advances of AI in agriculture, and by using X-ray images to monitor the quality of crambe seeds, the researchers of this study aim to evaluate the potency of deep learning (DL) models based on deep learning (CNNs). The proposed method described in this study uses seeds with various physical and physiological characteristics to build deep learning models. The models successfully distinguished seeds based on their vigour, germination, and internal tissue integrity with accuracy levels of 91, 95, and 82, respectively.

2.27 Deep Learning Approch for Seed Classification

(Loddo et al., 2021) introduced a deep learning approach for seed images classification, plus also gave a clue for retrieval of seed images. The researchers of this study have performed a study of more than 10 different CNN models, and applied classification accuracy and retrieval results of all those CNN architectures. The results of CNN classifiers are also investigated in this study and the most suitable CNN features that are required for seed image retrieval task are selected. The proposed model claims a finest accuracy of more then 95% on testing dataset, and the retrieval performance of proposed deep learning model is also analyzed, which yielded satisfying results. At the end, the researchers claim that the results obtained after applying CNNs to the seeds dataset can prove to be a best starting point for seed image classification and retrieval task, which in return would be much helpful for providing a strong support in agriculture and also in the fields of botany. After a brief go-through of this study, it could be given that performing deep learning & CNNs for both classification & retrieval turns out to be a fascinating application in crop seeds development.

(Xu et al., 2022) performed research on maize seeds dataset, which was based on deep learning (DL) and machine vision (MV), in which a classification model was developed for maize seed classification. This research has provided insights that maize comprises of one of the most essential crops in today's world that provides a means of utilization for humans. Traditional means for identification of maize seeds could be time-consuming and could also result in errors, so an intelligent support is needed for classification. The authors of this study suggest a quick image classification strategy for maize seeds that combines computer vision (CV) and deep learning (DL) to handle this challenging problem. 8080 maize seeds overall, divided into five classes, were gathered. The sample images were then divided into training and validation sets in an 80-20 ratio, and the data were improved. To identify and classify maize seeds, the suggested enhanced CNN network architecture, dubbed P-ResNet, was fine-tuned for transfer learning. The performance of the models is then contrasted. After application of multiple pre-trained CNN classifiers, such as AlexNet, GoogleNet, ResNet, MobileNet, etc., the accuracies of models came out to be more than 95% in research, while maximum accuracy among all these models was attained via P-ResNet. In addition, this research also states that by applying CNNs on maize seeds dataset, it can result in development of an effective model for classification.

2.28 Seeds Classification using Deep Learning Approach

(Gulzar et al., 2020) provided a contribution with the help of an intelligent system for seed classification with the help of convolsutional neural network (CNN). In this research, a CNN is applied on a dataset which comprises of 14 possible seeds by application of deep learning strategies. The approaches used in this study include hybrid weight modification, model task scheduling, and decaying learning rate. In this study, symmetry is used when sampling seed images to create data. When resizing and labelling the photos to extract their characteristics, the use of symmetry creates uniformity. As a result, the classification accuracy on training & testing sets were maximized, which are reported to be higher than the work already done in this field.

Paper	Headline	Accuracy (%)	Precision	Journal
Luo et al., (2021)	Classification of Weed Seeds Based on Visual Images & Deep Learning (DL)	98.44	-	Elsevier
De Medeiros et al., (2021)		89.3	-	Elsevier
Luddo et al., (2021)	Deep Learning Approach for Seed Image Classification & Retreival	96.56	-	Elsevier (Computer in Electronics & Agriculture)
Xu et al., (2022)	Research on Maize Seed Classification Based on Computer Vision & Machine Learning (ML)	96.44	-	MDPI
Gulzar et al., (2020)	CNN-based seed classification system	99	-	MDPI

Table 1: Comparison Table with other Researcher

Chapter 3

3.1 Research Methodology

The suggested research approach for conducting this study is that we will classify maize seeds using deep learning models. This method uses an intelligent system that performs intelligent seed classification using artificial intelligence (deep learning & CNNs). This system helps farmers and agricultural staff learn about a seed's development in advance so they can make the most of its cultivation. This section is divided into three phases: an input (data collecting) phase, a middle (data pre-processing) phase, and a functioning phase in which the deep learning model (CNN) for intelligent crop seed classification receives the data necessary to develop the intelligent system.

The loading of data relevant for intelligent categorization into the workspace as input is the initial part of the process. The dataset, which consists of photographs, may be utilised for classification to build an intelligent crop production system since the output of this model is a discrete value, i.e., seed classes. According to the proposed paradigm, data must be converted into numbers before being input into the system since computers can only understand numbers and cannot understand categories or string values. The next phase is the preprocessing phase, sometimes referred to as the intermediate or pre-processing phase in our nomenclature.

Input data is preprocessed using deep learning data preprocessing techniques during the preprocessing phase, also referred to as the intermediate phase. The removal of blank or missing data, converting category variables to numeric variables, and other pre-processing techniques are examples. The data will be cleaned and converted to numbers after the preprocessing stage, and these numbers may then be fed into a deep learning model, which will utilise them as input and learn data representations to make predictions about unseen situations. This step may include more complex data preparation techniques including sampling, under-sampling, and over-sampling.

Demonstrated below is a generalized deep learning model diagram, which is also applied for carrying on a deep learning model for crop seed classification in our research also.

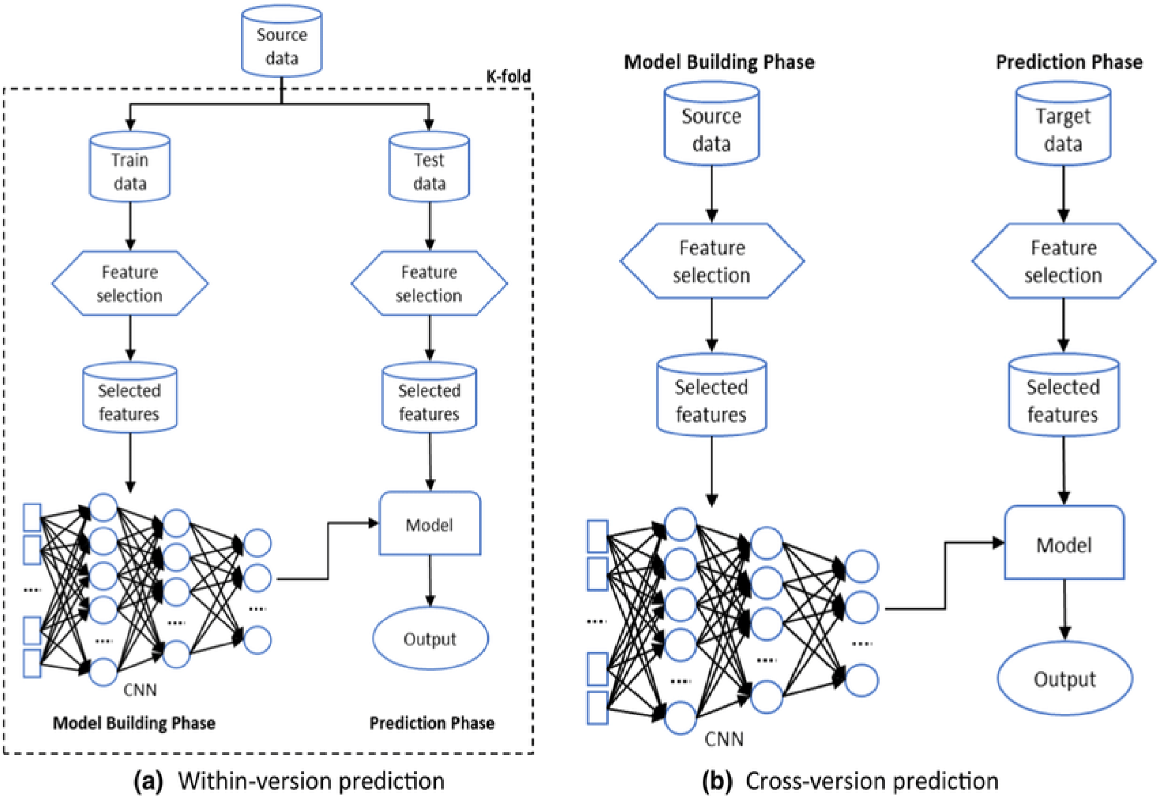

Figure 4: : Proposed Research Methodology for Seed Classification

The pre-processed data is submitted into a deep learning model (CNN) in the functional phase, which is the final stage, when the model learns image representation from the dataset. A loss function and an accuracy measure that are defined to determine a machine learning model's proficiency are used to train the representations. We present an intelligent system that learns from data and performs intelligent classification over unidentified or testing data because the proposed scheme is a classification model. In order to attain cutting-edge accuracy and low loss, the model will be trained to reduce categorical cross-entropy loss.

3.2 How Our Approach Differs

For dealing with image-based dataset, CNNs are most widely used for this task. Due to training time of a deep neural network and the complexity of problem, instead of manually developing a CNN from scratch, it's better to use pre-trained CNNs that are already trained on a much larger dataset, like ImageNet, which contains millions of images & 1000 classes. The model architectures can be found online, but to tine that model to a particular problem is the key challenge. In our case, we'll tune this model to crop classification image data, which contains 12 classes. Extracting useful features from images is the key task in CNNs, and pre-trained neural-nets have shown benchmark results on these tasks. In dealing with classification problem, along with accuracy score, it's also crucial to check the performance of model on each type of class. Our approach is based on computing model’s performance on each type of class (like 12 in our case) along with its accuracy, plus tuning a pre-trained model on our dataset.

Chapter 4:

4.1 Project Evaluation

After going through a thorough process of research methodology, literature, experimental evidences & background evaluation, now things are to be discussed which will decide how well our model is working on testing or validation dataset. This section will brief about the metrics that will be used for evaluating the skill or accuracy of proposed model on test dataset. Because the problem is a multi-class classification problem, in which we have to classify more then 2 kinds of maize seeds, therefore categorical cross-entropy will be used as loss metric for calculating loss, plus accuracy, precision, recall & F1 metrics will also be a part of model evaluation table. The target of this research is to develop a deep leaning model, preferably using a pre-trained CNN via transfer learning (TL) for boosting accuracy, making an intelligent classification system that would be able to classify the class of a maize seed to assist farmers in successful outlook of crop yield in farms.

4.2 Problem statement:

The quality of weed control is essential for the success of maize production. It is important to suppress weeds for the first six- to eight-week period after planting because during this time the crop faces intense competition from weeds for nutrients and water. Weed infections in cultivated crops cause annual output reductions. The amount of crop yield losses caused by weeds varies depending on the type of weed, the crop, and the surrounding environment. Generally speaking, crop losses might range from 10 to 100% depending on the extent of weed management used. One very rarely experiences zero yield loss from weeds. Plant involvement with the crop's development and growth leads to yield losses.

4.3 Algorithms used:

Since our problem is about image classification, transfer learning is used. A relevant pretrained deep learning model with convolutional layers is used for transfer learning:

- MobileNetV2

4.4 Dataset used:

There are 5,539 photos of weed and crop seedlings in this dataset. As shown in the photographs up top, the images are divided into 12 classes. In Danish agriculture, these classifications describe common plant species. Each class has rgb photos of plants in various stages of development.

The photos are in the png format and come in different sizes.

This dataset's V1 iteration was used in the Kaggle playground competition for plant species classification. This is version two. Multiple plants were present in some samples from V1. Now, such samples have been eliminated by the dataset's developers.

4.5 Transfer Learning:

In general, the term "learning transfer" refers to the process by which a model developed for one difficulty is used in some way for another similar problem. The Transfer Proposed method is a deep learning technique that trains a neural system model on a problem similar to the problem being solved. A new model is then trained on the relevant problem using one or more elements of the trained model.

Transfer learning has the advantage of reducing training time of neural network models and potentially reducing generalization error. The training procedure starts with previously used layer weights and can be modified to address new problems. In this usage, transfer learning is a type of weight initialization strategy. This can be an advantage if the first common error has much more tagged data than the topic of interest, as the closeness of the problem structure helps in both situations.

4.6 Implementation Techniques:

The paper describes the implementation of the model using Python Programming Language, by making use of following data science, machine learning and deep learning libraries:

- NumPy (python library for array computation)
- Pandas (python library for data preprocessing)
- Matplotlib (python library for data visualization)
- Seaborn (for data visualization)
- OS (python library for interacting with the operating system)

- Random (python library for generating random numbers)
- Counter (for counting the elements present in the container)
- CV2 (python library for computer vision, machine learning and image processing)
- Deep learning libraries for model building and evaluation.

4.7 Complete Flow:

Explained below is the complete flow, which will elaborate the complete methodology used to make this deep learning model.

- Importing the libraries
- Defining the folder paths
- Checking the distribution of the various classes for train, test and validation datasets
- Dealing with the imbalanced classes
- Building the model
- Adding the callbacks
- Compiling & fitting the model / Fine tuning
- Training testing accuracies & training testing loss graphs
  - Confusion matrix & classification report for each model

The complete flow of this paper has been mentioned above, and in below section we will be discussing each component of this paper step-by- step.

4.7.1 Importing The Necessary Python Libraries:

First, we will import all the important python machine learning and deep learning libraries for manipulating with our data e.g preprocessing, post processing, model building and finally model evaluation. Below attached is the code snippet screenshot.

```
import numpy as np
import pandas as pd
import os
import random
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
from collections import Counter
import cv2


from tensorflow.python.keras.applications.mobilenet_v2 import MobileNetV2, preprocess_input
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.callbacks import EarlyStopping, ReduceLROnPlateau


from tensorflow.keras.models import Model
from tensorflow.keras.layers import Dense, Flatten, Dropout, Input
```

Figure 5: Import Libarary in Python code

4.7.2 Defining the folder paths

We will make make a base directory and 3 subfolder directories named as train, test and validation and join these 3 directories together with the base directory using os.path.join. This will load all the images present in the train, test and validation directories. Below is the attached code snippet screenshot.

```
[ ]  base_dir = "/kaggle/input/kaggle-datasets/Seed_dataset"

     train_dir = os.path.join(base_dir,'train')
     val_dir = os.path.join(base_dir,'val')
     test_dir = os.path.join(base_dir, 'test')
```

Figure 6: Define Folder Path

4.7.3 Checking the distribution for train, test and validation directories:

We will check how many total no of classes are present inside the train, test and validation directories. In our case, we have total 6 classes each. The total no of images present inside train, test and validation directories are 1287, 412and 389. It also shows the class distribution for all 6 classes.

4.7.4 Dealing with imbalanced classes:

As our dataset is imbalnced, we will be using class weights for balancing our data. Below is the attached source code snippet.

```
train_datagen = ImageDataGenerator()
train_generator = train_datagen.flow_from_directory(train_dir)

counter = Counter(train_generator.classes)
max_val = float(max(counter.values()))
class_weights = {class_id : max_val/num_images for class_id, num_images in counter.items()}
print("\nThe class weights are : \n\n", class_weights)
```

Figure 7: Prepocess the Dataset

4.7.5 Buiding the model:

For model building, we will be setting the total no of classes to 6, the image size to (224,224) and the batch size to 32. Below is the attached source code snippet. In the next step, we will create the train, test and validation generators. These are called data generators. Data generators allow you to feed data into Keras in real-time while training the model. This way, we can make modifications to the data before feeding it to the neural network or even load it from the secondary memory. Below is the attached source code snippet.

```
[ ] # Create the Generators
    train_val_generator = ImageDataGenerator(
                                        preprocessing_function=preprocess_input,
                                        rotation_range=10,
                                        width_shift_range=0.2,
                                        height_shift_range=0.2,
                                        shear_range=0.1,
                                        zoom_range=0.2,
                                        horizontal_flip=True,
                                        vertical_flip=False,
                                        fill_mode='nearest'
                                        )

    # Train data generator
    train_data = train_val_generator.flow_from_directory(train_dir,
                                                target_size=IMAGE_SIZE,
                                                batch_size=BATCH_SIZE,
                                                class_mode='categorical',
                                                color_mode='rgb',
                                                shuffle=True
                                                )

    # Validation data generator
    val_data = train_val_generator.flow_from_directory(val_dir,
                                                target_size=IMAGE_SIZE,
                                                batch_size=BATCH_SIZE,
                                                class_mode='categorical',
                                                color_mode='rgb',
                                                shuffle=True
                                                )

    # Test data generator
    test_generator  = ImageDataGenerator(preprocessing_function=preprocess_input)
    test_data = test_generator.flow_from_directory(test_dir,
                                                target_size=IMAGE_SIZE,
                                                batch_size=1,
                                                class_mode='categorical',
                                                color_mode='rgb',
                                                shuffle=False
                                                )
```

Figure 8 Model Build

4.7.6 Adding the callbacks:

Use callbacks to perform actions during different stages of training. At the beginning or end of an epoch, before or after a single batch. This document uses stop Early and flatten lr as callback functions. If you quit early, your workout will end when your metrics stop working. "Lower

LR to plateau" will reduce the learning rate once the metric stops improving. Below is the attached source code snippet.

```
[ ]  reduce_lr = ReduceLROnPlateau(monitor='val_loss', factor=0.2, patience=10, min_lr=0.0001)
     early_stopping = EarlyStopping(monitor='val_loss', patience=10)
```

Figure 9: Adding Callbacks

4.7.7 Compiling & fitting the model:

Compile the model using Adam as the optimizer, categorical_cross entropy as the loss function, and accuracy as the scoring metric. Tune the model using the training and validation data, setting the number of epochs to 50 in this case, and calling the callback above. Fine tune the model by training the top layer and freezing the rest. In this case, we train the first 249 layers and freeze the rest. Recompile and retune the model. Below is the attached source code snippet.

```
NUM_CLASSES = 6
IMAGE_SIZE=[224, 224]
BATCH_SIZE=32
```

```python
from tensorflow.keras.applications.mobilenet_v2 import MobileNetV2
from tensorflow.keras.preprocessing import image
from tensorflow.keras.models import Model
from tensorflow.keras.layers import Dense, GlobalAveragePooling2D

# create the base pre-trained model
base_model = MobileNetV2(weights='imagenet', include_top=False)

# add a global spatial average pooling layer
x = base_model.output
x = GlobalAveragePooling2D()(x)
# let's add a fully-connected layer
x = Dense(1024, activation='relu')(x)
# and a logistic layer -- let's say we have 6 classes
predictions = Dense(6, activation='softmax')(x)

# this is the model we will train
model = Model(inputs=base_model.input, outputs=predictions)

# first: train only the top layers (which were randomly initialized)
# i.e. freeze all convolutional MobileNetV2 layers
for layer in base_model.layers:
    layer.trainable = False

# compile the model (should be done *after* setting layers to non-trainable)
# we will use adam
from tensorflow.keras.optimizers import Adam
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics = ['accuracy'])

# train the model on the new data for a few epochs
history = model.fit(train_data,
         validation_data = val_data,
         epochs = 50,
         callbacks = [reduce_lr, early_stopping])

# at this point, the top layers are well trained and we can start fine-tuning
# convolutional layers from mobilenet v2. We will freeze the bottom N layers
# and train the remaining top layers.

# let's visualize layer names and layer indices to see how many layers
# we should freeze:
for i, layer in enumerate(base_model.layers):
   print(i, layer.name)
```

```python
# we chose to train the top 2 mobilenet blocks, i.e. we will freeze
# the first 249 layers and unfreeze the rest:
for layer in model.layers[:249]:
   layer.trainable = False
for layer in model.layers[249:]:
   layer.trainable = True

# we need to recompile the model for these modifications to take effect
# we use Adam
from tensorflow.keras.optimizers import Adam
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics = ['accuracy'])

# we train our model again (this time fine-tuning the top 2 inception blocks
# alongside the top Dense layers
history = model.fit(train_data,
                    validation_data = val_data,
                    epochs = 50,
                    callbacks = [reduce_lr, early_stopping])
```

Figure 10: Compile and Fitting Model

4.7.8 Accuracy graph:

We will plot training accuracy vs testing accuracy graph for the above mentioned model.

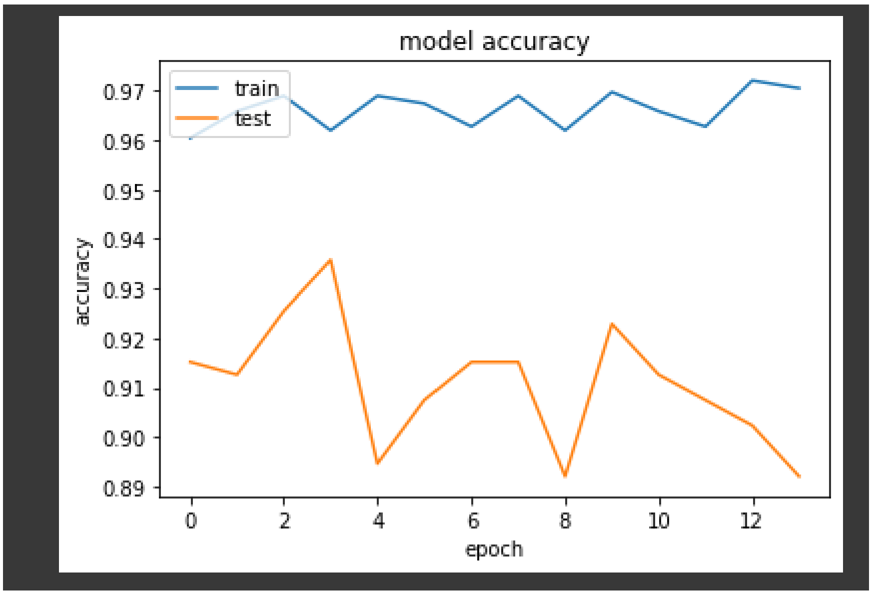

Figure 11: Epoc Vs Accuracy Graph

The graph indicates:

- The blue line represents training accuracy and indeed the orange line represents testing/validation accuracy.
- While the y-axis displays accuracy, the x-axis displays the number of epochs.
- The x-axis has a scale of 2, whereas the y-axis has a scale of 0.1.
- The graph shows that, with some changes and fluctuations in between epochs, validation accuracy follows training accuracy and reaches a maximum of 97.05% at the last epoch.

4.7.9 Loss graph:

We will plot training loss vs testing loss graph for the above mentioned model.

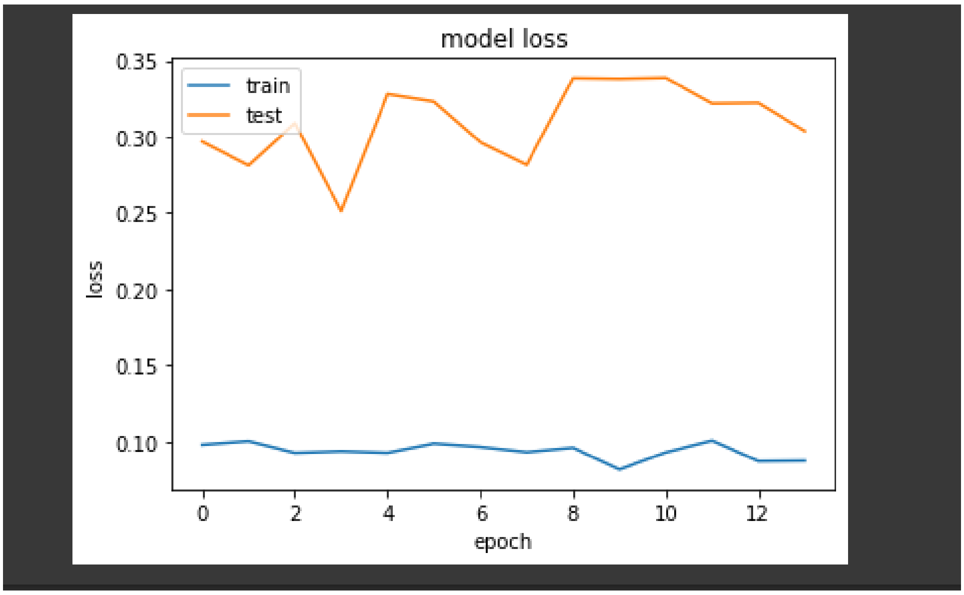

Figure 12: Epocs Vs. Loss Graph

The graph indicates that:

- The blue line represents the training loss, although the orange line represents the testing loss.
- The number of epochs is shown on the x-axis, while the model loss is shown on the y-axis.
- The x-axis has a scale of 2, whereas the y-axis has a scale of 0.1.
- The graph shows that the model loss decreases to 0.0873 during the last epoch while the validation loss follows training loss and fluctuates before ending up at 0.3037.
- The metric by which a machine learning or machine learning model learns current dataset is the loss function. The reduced loss indicates that our algorithm has gained knowledge through testing and training, which is doing well on both datasets.

Chapter 5

5.1 Conclusion

For countries depending more on agriculture, their economy can collapse and grow with increase or decrease in agricultural production, crop production & timely identification of crop seeds when they're growing at their initial stages. It could be much tricker for farmers and agricultural staff to timely figure out the class of a seed when it is growing in its initial stages, and manual identification could produce errors due to a chance of mistake by a human. Therefore, farmers and agriculture department can be easily assisted with the help of developments being carried out in Artificial Intelligence (AI) & Deep Learning (DL). These two fields have developed state-of-the-art algorithms like convolutional neural networks (CNNs) that are able to classify the given image into 2 or multiple classes with minimum loss. CNNs can be trained on images to produce high-quality results with a high accuracy. Therefore, for dealing with seeds of maize, we'll apply CNNs, pre-trained CNNs to be more precise, to perform maize seed classification into multiple classes with an accuracy that would be comparable with other researches in this area. This research will also highlight new challenges for upcoming researchers working in this field.

5.2 Future Work:

As in this study, 96.70 % accuracy on training dataset and 97.45 % accuracy on testing dataset was attained. To improve the accuracy and efficiently classify the category of seeds, Efficient Net B7 model will be used in future.

References

- Zhichkin, K., Nosov, V., Zhichkina, L., Dibrova, Z. and Cherepova, T., 2019, August. Development of evaluation model effectiveness of modern technologies in crop production. In *IOP Conference Series: Earth and Environmental Science* (Vol. 315, No. 2, p. 022023). IOP Publishing
- "Definition of CROP". *www.merriam-webster.com*. Retrieved June 20, 2017.
- Luo, T., Zhao, J., Gu, Y., Zhang, S., Qiao, X., Tian, W. and Han, Y., 2021. Classification of weed seeds based on visual images and deep learning. *Information Processing in Agriculture*.
- de Medeiros, A.D., Bernardes, R.C., da Silva, L.J., de Freitas, B.A.L., dos Santos Dias, D.C.F. and da Silva, C.B., 2021. Deep learning-based approach using X-ray images for classifying Crambe abyssinica seed quality. *Industrial Crops and Products*, *164*, p.113378.
- Loddo, A., Loddo, M. and Di Ruberto, C., 2021. A novel deep learning-based approach for seed image classification and retrieval. *Computers and Electronics in Agriculture*, *187*, p.106269.
- Xu, P., Tan, Q., Zhang, Y., Zha, X., Yang, S. and Yang, R., 2022. Research on Maize Seed Classification and Recognition Based on Machine Vision and Deep Learning. *Agriculture*, *12*(2), p.232.

- Gulzar, Y., Hamid, Y., Soomro, A.B., Alwan, A.A. and Journaux, L., 2020. A convolution neural network-based seed classification system. *Symmetry*, *12*(12), p.2018.

- Sabanci, K., Aslan, M.F., Ropelewska, E. and Unlersen, M.F., 2022. A convolutional neural network-based comparative study for pepper seed classification: Analysis of selected deep features with support vector machine. *Journal of Food Process Engineering*, *45*(6), p.e13955.
- Dönmez, E., 2022. Enhancing classification capacity of CNN models with deep feature selection and fusion: A case study on maize seed classification. *Data & Knowledge Engineering*, p.102075.
- Koklu, M. and Ozkan, I.A., 2020. Multiclass classification of dry beans using computer vision and machine learning techniques. *Computers and Electronics in Agriculture*, *174*, p.105507.
- Nie, P., Zhang, J., Feng, X., Yu, C. and He, Y., 2019. Classification of hybrid seeds using near-infrared hyperspectral imaging technology combined with deep learning. *Sensors and Actuators B: Chemical*, *296*, p.126630.
- Feng, L., Zhu, S., Liu, F., He, Y., Bao, Y. and Zhang, C., 2019. Hyperspectral imaging for seed quality and safety inspection: A review. *Plant methods*, *15*(1), pp.1-25.
- Singh, P., Nayyar, A., Singh, S. and Kaur, A., 2020. Classification of wheat seeds using image processing and fuzzy clustered random forest. *International Journal of Agricultural Resources, Governance and Ecology*, *16*(2), pp.123-156.
- Zhao, Y., Zhu, S., Zhang, C., Feng, X., Feng, L. and He, Y., 2018. Application of hyperspectral imaging and chemometrics for variety classification of maize seeds. *Rsc Advances*, *8*(3), pp.1337-1345.
- Xia, Y., Xu, Y., Li, J., Zhang, C. and Fan, S., 2019. Recent advances in emerging techniques for non-destructive detection of seed viability: A review. *Artificial Intelligence in Agriculture*, *1*, pp.35-47.
- Fazel-Niari, Z., Afkari-Sayyah, A.H., Abbaspour-Gilandeh, Y., Herrera-Miranda, I., Hernández-Hernández, J.L. and Hernández-Hernández, M., 2022. Quality Assessment of Components of Wheat Seed Using Different Classifications Models. *Applied Sciences*, *12*(9), p.4133.
- Lv, X., Ming, D., Chen, Y. and Wang, M., 2019. Very high resolution remote sensing image classification with SEEDS-CNN and scale effect analysis for superpixel CNN classification. *International Journal of Remote Sensing*, *40*(2), pp.506-531.
- Qiu, Z., Chen, J., Zhao, Y., Zhu, S., He, Y. and Zhang, C., 2018. Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network. *Applied Sciences*, *8*(2), p.212.

- Bagheri, A. and Heydari, Z., 2021. Classification of weed seeds using image processing and pattern recognition methods. *Iranian Journal of Weed Science*, *17*(2), pp.65-77.
- F Shalabi, L., Hamdy, N. and Hussein, E., 2022. Taxonomic significance of seed characters and SDS-PAGE analysis in the classification of Ericaceae. *Taeckholmia*, *42*(1), pp.1-11.
- Altuntaş, Y., Cömert, Z. and Kocamaz, A.F., 2019. Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach. *Computers and Electronics in Agriculture*, *163*, p.104874.
- Uzal, L.C., Grinblat, G.L., Namías, R., Larese, M.G., Bianchi, J.S., Morandi, E.N. and Granitto, P.M., 2018. Seed-per-pod estimation for plant breeding using deep learning. *Computers and electronics in agriculture*, *150,* pp.196-204.
- Tang, Y., Wang, X., Harrison, A.P., Lu, L., Xiao, J. and Summers, R.M., 2018, September. Attention-guided curriculum learning for weakly supervised classification and localization of thoracic diseases on chest radiographs. In *International Workshop on Machine Learning in Medical Imaging* (pp. 249-258). Springer, Cham.
- Singh, K., Saccon, F. and Joseph, D., 2022. Incremental two-network approach to develop a purity analyzer system for canola seeds. *Electronic Imaging*, *34*, pp.1-7.
- Jiaoyang, Z. and Xingshi, H., 2019, October. The Application of Support Vector Machine with Elastic Net Regularizationin in Classification of Wheat Seeds. In *Journal of Physics: Conference Series* (Vol. 1325, No. 1, p. 012036). IOP Publishing.
- Matera, T.C., Pereira, L.C., Braccini, A.L., Krzyzanowski, F.C., Scapim, C.A., Piana, S.C., Marteli, D.C.V., Pereira, R.C., Ferri, G.C. and Suzukawa, A.K., 2019. Accelerated aging test and its relationship to physiological potential of soybean seeds. *Journal of Seed Science*, *41*, pp.301-308.
- Singh Thakur, P., Tiwari, B., Kumar, A., Gedam, B., Bhatia, V., Krejcar, O., Dobrovolny, M., Nebhen, J. and Prakash, S., 2022. Deep transfer learning based photonics sensor for assessment of seed-quality.
- ElMasry, G., Mandour, N., Al-Rejaie, S., Belin, E. and Rousseau, D., 2019. Recent applications of multispectral imaging in seed phenotyping and quality monitoring—An overview. *Sensors*, *19*(5), p.1090.
- Fu, L., Sun, J., Wang, S., Xu, M., Yao, K., Cao, Y. and Tang, N., 2022. Identification of maize seed varieties based on stacked sparse autoencoder and near-infrared hyperspectral imaging technology. *Journal of Food Process Engineering*, p.e14120.
- Ramírez, N. and Briceño, H., 2022. Ecology of morphological fruit types, and fruit and seed colors in 27 Venezuelan plant communities. *Acta Oecologica*, *116*, p.103838.

- Cervantes, E. and Martin Gomez, J.J., 2018. Seed shape quantification in the order Cucurbitales.
- Guerin, C., Serret, J., Montúfar, R., Vaissayre, V., Bastos-Siqueira, A., Durand-Gasselin, T., Tregear, J., Morcillo, F. and Dussert, S., 2020. Palm seed and fruit lipid composition: phylogenetic and ecological perspectives. *Annals of botany*, *125*(1), pp.157-172.
- Hasan, H.M., 2022. Study the Effect of Different Polar Extracts of Parsley Seeds (Petroselinum Crispum) and Its Phenolic Compounds as Antioxidants. *Annals of the Romanian Society for Cell Biology*, *26*(01), pp.1898-1905.
- Yuan-Yuan, Y.U.A.N., Jia-Lin, D.U., Zhao-Peng, G.E.N.G., Yi, L.I., Xu, L.U., Yue, L.I.U. and Jun-Li, W.A.N.G., 2022. Exploring the composition of Syringa reticulata subsp. amurensis seed and its underlying mechanism against chronic bronchitis. *Chinese Journal of Analytical Chemistry*, *50*(9), p.100132.
- Wakholi, C., Kandpal, L.M., Lee, H., Bae, H., Park, E., Kim, M.S., Mo, C., Lee, W.H. and Cho, B.K., 2018. Rapid assessment of corn seed viability using short wave infrared line-scan hyperspectral imaging and chemometrics. *Sensors and Actuators B: Chemical*, *255*, pp.498-507.
- Pereira, D.F., Bugatti, P.H., Lopes, F.M., de Souza, A.L.S.M. and Saito, P.T.M., 2020. Assessing Active Learning Strategies to Improve the Quality Control of the Soybean Seed Vigor. *IEEE Transactions on Industrial Electronics*, *68*(2), pp.1675-1683.
- Martín Gómez, J.J., Gutiérrez del Pozo, D. and Cervantes, E., 2019. Seed shape quantification in the Malvaceae reveals cardioid-shaped seeds predominantly in herbs.
- Torres-Cortés, G., Bonneau, S., Bouchez, O., Genthon, C., Briand, M., Jacques, M.A. and Barret, M., 2018. Functional microbial features driving community assembly during seed germination and emergence. Frontiers in plant science, 9, p.902.
- Coysh, T. and Mead, S., 2022. The Future of Seed Amplification Assays and Clinical Trials. Frontiers in Aging Neuroscience, 14.
- Kurmi, Y. and Gangwar, S., 2022. A leaf image localization based algorithm for different crops disease classification. *Information Processing in Agriculture*, *9*(3), pp.456-474.
- Amézquita, E.J., Quigley, M.Y., Ophelders, T., Landis, J.B., Koenig, D., Munch, E. and Chitwood, D.H., 2022. Measuring hidden phenotype: Quantifying the shape of barley seeds using the Euler Characteristic Transform. *in silico Plants*, *4*(1), p.diab033.
- Le, Q.T. and Sritontip, C., 2021. Effects of nano-bubbles on seed germination of muskmelon. *Vietnam Journal of Science, Technology and Engineering*, *63*(3).

- Khazaei, H., Subedi, M., Nickerson, M., Martínez-Villaluenga, C., Frias, J. and Vandenberg, A., 2019. Seed protein of lentils: Current status, progress, and food applications. *Foods*, *8*(9), p.391.
- Polder, G., Blok, P.M., De Villiers, H.A., Van der Wolf, J.M. and Kamp, J., 2019. Potato virus Y detection in seed potatoes using deep learning on hyperspectral images. Frontiers in plant science, 10, p.209.
- Zhu, S., Zhou, L., Gao, P., Bao, Y., He, Y. and Feng, L., 2019. Near-infrared hyperspectral imaging combined with deep learning to identify cotton seed varieties. Molecules, 24(18), p.3268.
- Nie, P., Zhang, J., Feng, X., Yu, C. and He, Y., 2019. Classification of hybrid seeds using near-infrared hyperspectral imaging technology combined with deep learning. Sensors and Actuators B: Chemical, 296, p.126630.
- Wu, N., Liu, F., Meng, F., Li, M., Zhang, C. and He, Y., 2021. Rapid and accurate varieties classification of different crop seeds under sample-limited condition based on hyperspectral imaging and deep transfer learning. Frontiers in bioengineering and biotechnology, 9.
- Khed, V., Jaleta, M. and Krishna, V., 2021. Seed Delivery Pathways and Farmers' Access to Improved Wheat Varieties in Ethiopia and India.
- Picon, A., Seitz, M., Alvarez-Gila, A., Mohnke, P., Ortiz-Barredo, A. and Echazarra, J., 2019. Crop conditional Convolutional Neural Networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions. *Computers and Electronics in Agriculture*, *167,* p.105093.
- Toda, Y., Okura, F., Ito, J., Okada, S., Kinoshita, T., Tsuji, H. and Saisho, D., 2020. Training instance segmentation neural network with synthetic datasets for crop seed phenotyping. *Communications biology*, *3*(1), pp.1-12.

Ethics From

Ethical clearance for research and innovation projects

Project status

Status

Approved

Actions

Date	Who	Action	Comments
09:34:00 09 September 2022	Femi Isiaq	Supervisor approved	
09:33:00 09 September 2022	Bhumilkumar Patel	Principal investigator submitted	

Get Help