IMAGE PROCESSING-BASED TRANSFER LEARNING APPROACH FOR CLASSIFICATION OF VEGETABLE SEEDS OF BANGLADESH

BY

MD. ATIQUE SHAHRIAR
ID: 203-15-14510

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

MD. SAZZADUR AHAMED
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

ABDUS SATTAR
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled "Image processing-based transfer learning approach for classification of vegetable seeds of Bangladesh", submitted by Md. Atique Shahriar to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 15th July, 2024.

BOARD OF EXAMINERS

Taimur Ahad

Dr. Md. Taimur Ahad (MTA) — **Board Chairman**
Associate Professor & Associate Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Shayla Sharmin

Ms. Shayla Sharmin (SS) — **Internal Examiner 1**
Sr. Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Mayen Uddin Mojumdar (MUM) — **Internal Examiner 2**
Sr. Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

M Islam

Dr. Md. Manowarul Islam — **External Examiner**
Associate Professor
Department of CSE
Jagannath University

DECLARATION

I hereby declare that this project has been done by Md. Atique Shahriar under the supervision of **Md. Sazzadur Ahamed, Assistant Professor, Department of Computer Science and Engineering, Daffodil International University**. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Md. Sazzadur Ahamed
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Atique

Md. Atique Shahriar
ID: 203-15-14510
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty for His divine blessing making it possible for me to complete the final year project successfully.

I am grateful and wish our profound indebtedness to **Md. Sazzadur Ahamed, Assistant Professor,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of "*Computer Vision*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to the **Head of the Department of CSE**, for his kind help in finishing this project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

Research on the agriculture sector has increased within the past few years. This study focuses on developing a transfer learning-based approach to improve vegetable seed recognition and verification, addressing the crucial demand for accurate seed classification. The agricultural economy of Bangladesh heavily relies on the quality of seeds, yet manual classification methods are time-consuming and error-prone. There are several features to classify a seed, such as color, length, texture, size, shape etc. This research evaluates the classification performance of three transfer learning models MobileNetV2, Inception V3, and NASNetMobile on a dataset comprising 6,000 preprocessed images, with each of the 12 vegetable seed varieties represented by 500 images. The research focuses on enhancing model performance through adjustments in image resolution and epoch parameters. Out of all the experiments, MobileNetV2 provides an accuracy rate of 99.50% and a loss of 1.57%, this performance is validated by the test dataset. These implications enhance agricultural seed classification by improving efficiency and accuracy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AI	Artificial Intelligence
ML	Machine Learning
CNN	Convolutional Neural Network
SVM	Support Vector Machine
KNN	K-Nearest Neighbor
DT	Decision Tree
MLP	Multilayer Perceptron
NB	Naïve Bayes
LR	Linear Regression
ReLU	Rectified Linear Unit
CPU	Central Processing Unit
RAM	Random Access Memory
SSD	Solid State Drive
GPU	Graphics Processing Unit
BADC	Bangladesh Agricultural Development Corporation

CHAPTER 1
Introduction

1.1 Overview

This study explores the pivotal role of seed quality in enhancing agricultural productivity in Bangladesh, focusing on the development of an automated vegetable seed classification system using transfer learning. Given the significant impact of seed quality on crop yields and the challenges posed by the scarcity of high-quality seeds, this research aims to streamline and enhance the seed identification process. By leveraging advanced machine learning techniques, the study seeks to provide an efficient, accurate, and reliable solution for classifying diverse vegetable seed varieties. The methodologies employed include data collection, model training, and validation, while the outcomes are evaluated in terms of accuracy, efficiency, and practical applicability in the agricultural sector of Bangladesh.

1.2 Background and Present State

The economic development of Bangladesh is significantly influenced by its agricultural sector, which is important to the nation's overall growth. Within agriculture, the optimization of crop yield and quality significantly depends upon various factors, foremost among them being the quality of seeds [1]. Seeds are the fundamental organic ingredient by which plants reproduce their species, making them essential for farming [2]. Given that the vast majority of plants grow from seeds, they represent the most valuable resource for farming populations. Consequently, the full benefit of such inputs and agricultural procedures cannot be realized if high-quality seed is not employed. Healthy seeds are the fundamentally important input that all other inputs depend on for their full effectiveness. Notably, the private sector plays a substantial role in vegetable seed production, with approximately 95% of vegetable seeds being produced by private entities, whereas the government sector, such as BADC, accounts for only a minor portion [1].

Though seed quality has a significant impact on crop yield, the scarcity of high-quality seeds hampers farmers' ability to improve food production. Bangladesh boasts a diverse array of crops, including paddy, wheat, oilseeds, jute, fruits (such as mango, jackfruit, litchi, and coconut), vegetables (like tomato, cabbage, cucumber, pumpkin, and gourd), and spices (including coriander, garlic, ginger, and onion) [3]. With close to 50 percent of the population of Bangladesh engaged in agriculture and over 70 percent of the land dedicated to crop cultivation [4]. Identifying seed varieties poses a significant challenge from various types. Different seed varieties have different color, size, and shape features, which makes identification more difficult. To solve this problem, seed classification is therefore very important. Classification enables individuals to easily identify their desired seed variety by assessing its shape and skin consistency, thereby reducing the risk of deception when purchasing seeds from local markets.

Furthermore, seed classification not only aids experienced farmers as well as benefits younger individuals with useful knowledge for future activities. The capacity to differentiate between various seed varieties becomes more and more crucial with the developing trend of urban agriculture, where many people in cities cultivate fruits and vegetables in their corridors and on the roofs of their homes. There is still a shortage of research on the classification of mixed vegetable seeds, despite the fact that many scholars have focused on classifying individual crop seeds like rice, soyabean oil, wheat, dry beans etc. Obtaining a trustworthy dataset for this kind of classification is extremely challenging because high-quality seed sources are difficult to obtain.

1.3 Problem Statement

Manual seed classification is time-consuming and error-prone, hindering the optimization of crop yield and quality. The scarcity of high-quality seeds further challenges farmers' ability to improve food production in Bangladesh. The challenge of identifying seeds is made more difficult by a shortage of studies on the classification of seeds from vegetable crops.

1.4 Objectives

The project aims to address the need to improve seed recognition and verification within Bangladesh's agricultural sector by developing a transfer learning-based approach to effectively identify and classify vegetable seeds. The manual process of seed classification is not only time-consuming but also prone to inaccuracies, which can hamper the effectiveness of agricultural practices. By leveraging automation through transfer learning, the project significantly improves the efficiency and reliability of seed classification tasks. This approach holds the ability to completely change seed identification processes by providing Bangladeshi farmers and agricultural professionals with a precise and efficient solution to their difficulties.

1.5 Scope and Limitations

The project scope includes the development of a transfer learning-based approach for accurately classifying vegetable seeds. Additionally, the scope includes tackling challenges associated with seed variety, quality, and dataset availability, aiming to provide comprehensive solutions to enhance seed recognition within the agricultural sector.

1.6 Report Organization

The thesis is organized into seven main chapters:

Chapter 1 – Introduction: This introductory chapter provides an overview of the project's, background and present state, problem statement, objectives, scope and limitations.

Chapter 2 – Literature Review: Discusses existing research in the domain, summarizing findings, limitations, and challenges encountered in this chapter.

Chapter 3 – Methodology: The research methodology chapter outlines the methodology to be employed, including hardware/software requirements, project management and financial analysis.

Chapter 4 – Implementation: The chapter outlines the data collection procedures, model details and details of evaluation matrices for findings models performance.

Chapter 5 – Result and Analysis: Presents the experimental results, analysis of findings, and discussions based on the results.

Chapter 6 – Impact on Society, Environment and Sustainability: Discusses the societal impacts, environmental implications, ethical aspects, and sustainability plans arising from the research.

Chapter 7 – Conclusion and Future Work: Summarizes the study's findings, draws conclusions, offers recommendations, suggests directions for future research and limitations.

1.7 Summary

The economic development of Bangladesh relies heavily on agriculture, where seed quality is paramount. Manual seed classification is time-consuming, exacerbated by a scarcity of studies on vegetable seeds. This project aims to improve seed recognition by developing a transfer learning-based approach. Automation can revolutionize seed identification, benefiting both experienced farmers and urban agriculturists. The project's scope includes addressing problems with dataset availability and vegetable seed variety, as well as offering complete solutions for seed identification in the agriculture industry.

CHAPTER 2

Literature Review

2.1 Overview

In recent years, significant advancements have been made in seed classification and categorization research. Researchers have employed various methodologies, including deep learning, machine learning, and image processing techniques, to classify different types of seeds accurately. This section provides a comprehensive summary of recent studies, highlighting their methodologies, achieved accuracies, key findings, challenges, and future directions in seed classification research.

2.2 Related Works

In 2019, Zhu et al. utilized transfer learning for identifying soybean seed varieties [5]. They trained their model on 9600 seeds from ten varieties, augmenting the dataset to enhance diversity, with the primary collection comprising 1200 images. They employed image rotation techniques such as 90°, 180°, and 270° rotations for augmentation. Following fine-tuning, they achieved a notable accuracy of 97.2% using the NASNetLarge pretrained model.

In the same year [6], using computer vision and machine learning techniques Murat Koklu and Ilker Ali Ozkan developed a multiclass classification system to recognize seven different varieties of dry beans, comprising a dataset of 13,611 grains. Utilizing segmentation and feature extraction methods, they extracted a total of 16 features. Subsequently, various machine learning models including Multilayer Perceptron (MLP), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree (DT) were trained using 10-fold cross-validation. Among these models, SVM provides the highest accuracy at 93.13%.

In addition, Md Mushfiqur Rahman and Pratik Barua in 2021 proposed an ensemble approach for fruit identification, employing four pre-trained CNN models (VGG-16, InceptionV3, Xception, and ResNet) on 20 different fruit seeds [7]. Ensemble learning is a popular approach because it improves forecasting accuracy and resilience by aggregating predictions from various models, reducing errors and biases for better performance [8]. 88% accuracy was achieved by using ensemble approach after combining the results of each individual model.

Furthermore, utilizing an ensemble learning approach Khatri et al. in 2022 also developed a seed classification system on wheat seeds [9]. Despite the limited dataset containing 210 samples from the UCI online platform, with each variety comprising 70 images, the researchers achieved a remarkable accuracy of 95% through ensemble classifiers employing hard voting.

Utilizing a vast dataset comprising 14 rice varieties, each with 3,500 seed samples totaling close to 50,000 seeds Kiratiratanapruk et al. in 2020 developed an automatic grading machine of paddy rice seed classification using machine learning [10]. A screening method detected unusual physical seed samples. Five pre-trained deep learning models and four statistical machine learning techniques were used to compare the classification performance. SVM achieved 90.61% accuracy, while InceptionResNetV2 reached 95.15%.

Introduced a custom classification model [11] with 16 layers Bulent TUGRUL in 2022 for rice seeds that's grown in Turkey. Despite customizing a model, the highest accuracy of 97% was attained using VGG, while the custom model yielded 96% accuracy. This evaluation was conducted on a dataset of 6833 images resized to 66x64x3 pixels.

In 2020, Keya et al. introduced a plant seed classification system based on deep learning utilizing Convolutional Neural Networks (CNNs) [12]. They gathered images of five types of plant seeds, totaling 1250 photos, which were then resized to 128x128 pixels to speed up processing. Through the application of CNNs, the researchers attained an accuracy of 89% in training and 93% in validation phases.

Gulzar et al. in 2018, proposed CNN based seed classification system [13] with 14 commonly known seeds, such as black pepper, coriander, cumin, onion etc, utilizing a dataset of 2733 images. They employed rescaling to reduce the image size to 224x224x3

without altering the shape, and augmented the dataset to generate additional training images. The researchers fine-tuned the VGG16 model, substituting the last five layers with the average pooling layer, the flatten layer, the dense layer, the dropout layer and the softmax layer, achieving an impressive accuracy of 99%.

In the same year, A. V. Seethalakshmi and H. Salome Hemachitra introduced a system that identifies complex type seed varieties using image processing techniques [14]. They used Ant Colony Optimization (ACO) strategy to utilize feature selection and Support Vector Machine (SVM) is employed for classification which provides about 93% accuracy during prediction.

In 2021, Pakeerathan et al. presented a study on the classification of vegetable plant pests using deep transfer learning [15]. They created a dataset comprising 500 vegetable pest image samples across ten categories. Employing two pre-trained CNN models, VGG16 and InceptionV3, originally trained on the ImageNet dataset, they fine-tuned the models with the limited number of pest image samples. Remarkably, utilizing the InceptionV3 model led to an accuracy of 99.8%.

In 2022, Jaithavil et al. [16] employed a system for classification of paddy seeds utilizing deep learning techniques. They made a custom dataset using 1200 images and evaluated three deep learning models VGG16, InceptionV3, MobileNetV2. Among these models, InceptionV3 exhibited the highest accuracy 83.33% and the lowest test loss of 28.41%.

A machine learning model for classifying wheat into fresh and rotting categories was developed in 2023 by Agarwal et al [17]. Their dataset consisted of 400 photos, evenly distributed across fresh and decaying wheat samples. Shadows in the photos were removed during preprocessing. The SVM classifier provided the highest performance; it was able to distinguish between fresh and rotting samples based on color features with an outstanding 93% accuracy.

2.3 Comparison between existing works

Although there have been significant improvements in seed classification, there are notable differences in methodology, dataset sizes, preprocessing techniques, and achieved accuracies across all the studies. While some studies find a broader classification, others concentrate on specific seed types. However, there is broad acceptance on the efficiency of deep learning and machine learning algorithms in achieving high accuracies across numerous seed classification tasks. The accuracy achieved in several studies by other authors is shown in Table 2.3.1.

TABLE 2.3.1. COMPARATIVE ANALYSIS OF PREVIOUS WORK

Author Name	**Used Algorithm**	**Best Accuracy with Algorithm**
Zhu et al. [5]	AlexNet, ResNet18, Xception, InceptionV3, DenseNet201, and NASNetLarge	NASNetLarge = 97.2%
Murat Koklu and Ilker Ali Ozkan [6]	MLP, SVM, KNN, and DT	SVM= 93.13%
Md Mushfiqur Rahman and Pratik Barua [7]	VGG-16, InceptionV3, Xception, and ResNet	Ensemble = 88%
Khatri et al. [9]	KNN, DT, and NB	Ensemble = 95%
Kiratiratanapruk et al. [10]	LR, LDA, KNN, SVM, VGG16, VGG19, Xception, InceptionV3, and InceptionResNetV2	InceptionResNetV2 = 95.15%
Bulent TUGRUL [11]	VGG, Custom CNN, ResNet, EfficientNet.	VGG = 97%
Keya et al. [12]	CNN	CNN=93%
Gulzar et al. [13]	VGG16	VGG16 = 99%
A. V. Seethalakshmi and H. Salome Hemachitra [14]	SVM	SVM = 93.48%
Pakeerathan et al. [15]	VGG16 and InceptionV3	InceptionV3 = 99.8%
Jaithavil et al. [16]	VGG16, InceptionV3, MobileNetV2	InceptionV3 = 83.33%
Agarwal et al [17]	SVM, KNN, MLP, NB	SVM = 93%

2.4 Open Issues

This research addresses several key issues in the classification of mixed vegetable seeds, an area that is less explored compared to crops like rice, wheat, and soybean. The main problems include the lack of custom and diverse datasets necessary for accurate seed classification. Current research has achieved high accuracy for specific seeds but falls short for mixed vegetable seeds. This is due to limited research in the vegetable seed sector.

Additionally, there is a need for standardized evaluation metrics to consistently assess model performance. The small size and poor quality of available datasets also hinder the development of robust classification models. Current methods struggle with the variability in seed characteristics such as color, size, texture, and shape.

To address these issues, this study will focus on:

- Standardizing evaluation metrics for consistent model assessment.
- Increasing the size and diversity of datasets for better training and validation.
- Exploring transfer learning approaches to improve accuracy in vegetable seeds classification.

By solving these problems, the research aims to develop more reliable and accurate methods for vegetable seed classification, ultimately helping farmers improve crop yields and quality.

2.5 Summary

Seed classification research has progressed because of the use of numerous techniques such as deep learning, transfer learning, and ensemble learning. Even while some research has produced results with high accuracy for particular seed types, there are still challenges related to dataset variability and standardization of evaluation metrics. In order to create vegetable seed classification techniques that are more reliable and broadly applicable.

CHAPTER 3
Methodology

3.1 Overview

In order to classify vegetable seeds accurately, an efficient approach is necessary. To avoid biases, this involves gathering a significant set of seed images from a variety of vegetable categories. The images are then resized for computing efficiency, and important features are extracted using pre-trained CNNs. The effectiveness of the model can be determined by metrics like accuracy, which are utilized after model selection, training, and evaluation.

3.2 Proposed Methodology

This study explores the exact classification of vegetable seeds through the use of modern artificial intelligence (AI) techniques. This study focuses on the application of transfer learning which utilizes knowledge from models that have already been trained on vast data sets to perform seed classification tasks. The aim is to enhance the accuracy and efficiency of identifying between different kinds of vegetable seeds, which is crucial for various agricultural applications.

Traditional methods for classifying seeds are frequently error-prone and labor-intensive. This work proposes to address these limitations by combining deep Convolutional Neural Networks (CNNs) with transfer learning. Because deep CNNs can extract complex patterns and characteristics from visual input, they are especially well-suited for image analysis.

The study involves a comparative analysis of several predefined models, including MobileNetV2, Inception V3, and NASNetMobile. These models are selected based on their proven effectiveness in image classification tasks. To find the best configuration for seed classification, each model is evaluated across a range of cases, including varying image resolutions and epoch settings. Fig. 3.2.1, illustrates the workflow of the suggested model.

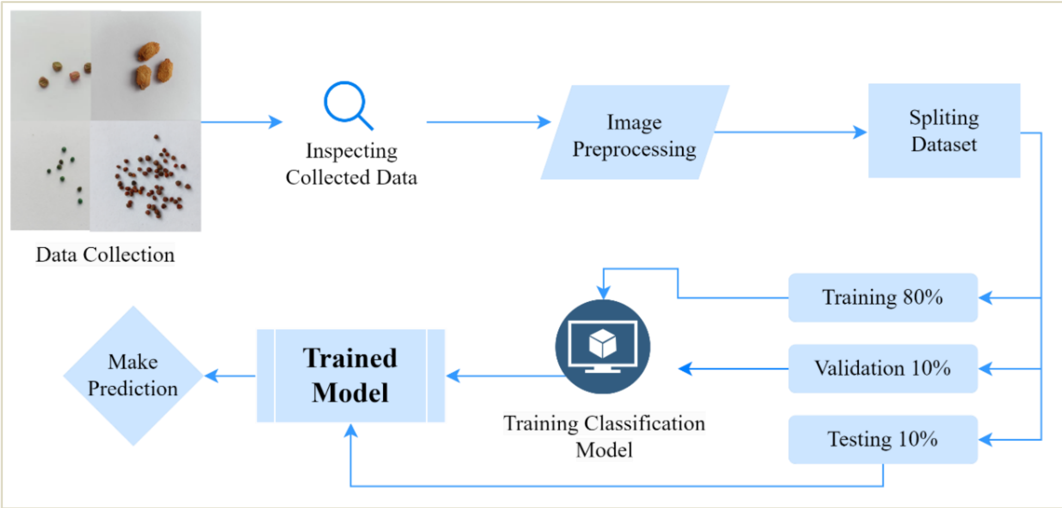

Figure 3.2.1. Workflow of this study.

In overall, the study area combines AI, deep learning, and agricultural science to provide a reliable and efficient system for classifying vegetable seeds. The findings are expected to provide valuable insights that can improve seed sorting and selection processes, contributing to better crop management and higher agricultural productivity.

3.3 Hardware/Software Requirements

The successful implementation of seed classification depends on meeting specific hardware and software requirements. These elements are crucial for ensuring efficient model training and seamless system operation.

Hardware Requirements: A high-performance local machine with an Intel i7 11th Gen CPU, 16 GB RAM, and a 512 GB SSD supports efficient processing. Model training requires substantial graphical processing power. Utilizing GPUs accelerates the training process, making them the preferred choice for handling complex models. Additionally, sufficient RAM is necessary to manage the dataset and the computational demands of the model.

Software Requirements: The primary coding environment for this research is Google Colab. Each model, including MobileNetV2, Inception V3, and NASNetMobile, was trained in the cloud using Google Colab's resources with a T4 GPU. For resizing images locally, Visual Studio Code is utilized. Additionally, popular deep learning framework TensorFlow is employed for the development, training, and evaluation of the CNN models. These frameworks provide essential libraries and tools to facilitate the creation and optimization of deep learning models.

By fulfilling these hardware and software requirements and optimizing the model architecture, the study aims to develop an effective system for automated seed classification.

3.4 Project Management and Financial Analysis

The data collection method includes several key steps designed to ensure the acquisition of high-quality seeds. First, comprehensive online research is carried out to identify trustworthy sources known for providing high-quality seeds. Furthermore, consultations with experts in numerous nurseries are undertaken to gather insights and recommendations. After identifying expected sources, go to the seed market and personally gather the seeds. Table 3.4.1 shows the estimated cost for this project.

TABLE 3.4.1. ESTIMATED COST FOR THE PROJECT.

SN	Components	Estimated Cost (BDT)
01.	Purchasing seeds from 12 different categories of vegetables	1600
02.	Documentation	1000
03.	Contingency	400
Total Estimated Cost		**3000**

3.5 Summary

This study investigates the classification of vegetable seeds using advanced AI techniques, focusing on transfer learning to enhance accuracy and efficiency. Traditional methods are often error-prone and labor-intensive; this research leverages deep Convolutional Neural Networks (CNNs) for superior image analysis. It compares models like MobileNetV2, Inception V3, and NASNetMobile to determine the best configuration for seed classification. The study involves collecting and preprocessing high-resolution images of twelve seed categories. The findings aim to improve seed identification processes, benefiting agricultural productivity in Bangladesh.

CHAPTER 4
Implementation

4.1 Overview

This study focuses to revolutionize agricultural practices in Bangladesh through advanced AI techniques aimed at enhancing vegetable seeds classification. By leveraging transfer learning with MobileNetV2, InceptionV3, and NASNetMobile models, the research addresses the critical need for accurate seed quality assessment. A dataset of 6000 high-resolution images across twelve seed categories is meticulously processed to ensure data uniformity and clarity. Comprehensive evaluations utilizing accuracy, precision, recall, and F1 score metrics provide robust insights into the models' performance, highlighting MobileNetV2 as the most efficient choice for seed classification.

4.2 Data Collection Procedure

A. Dataset Collection:

Collecting better quality images is a challenge for this seed classification problem. Twelve categories such as Bitter gourd, Bottle gourd, Peas, Beans, Red amaranth, Cabbage, Cauliflower, Carrot, Cucumber, Capsicum, Bean, Broccoli, Brinjal are gathered. These images are taken physically via smartphone cameras, named the Realme 8 with a 64-megapixel camera and the iPhone 11 with a 12-megapixel camera. That device provides a better resolution (3468*3468 and 3024*3024 respectively) for each image. The collection process involved acquiring seeds from seed markets.

TABLE 4.2.1. NUMBER OF TOTAL COLLECTED DATA.

Category	Quantity before preprocessing
Bitter Gourd	588
Bottle Gourd	582
Cucumber	544
Capsicum	556

Peas	542
Beans	571
Brinjal	553
Carrot	544
Broccoli	567
Cabbage	533
Cauliflower	557
Red Amaranth	531
Total	**6668**

Before being captured on camera, the seeds were placed on a white background in the proper daylight. Approximately 6668 images have been collected, covering twelve categories of vegetable seeds, with each category containing around 550 images to ensure a balanced dataset. Fig. 4.2.1. and Fig. 4.2.2. represents the sample images of dataset.

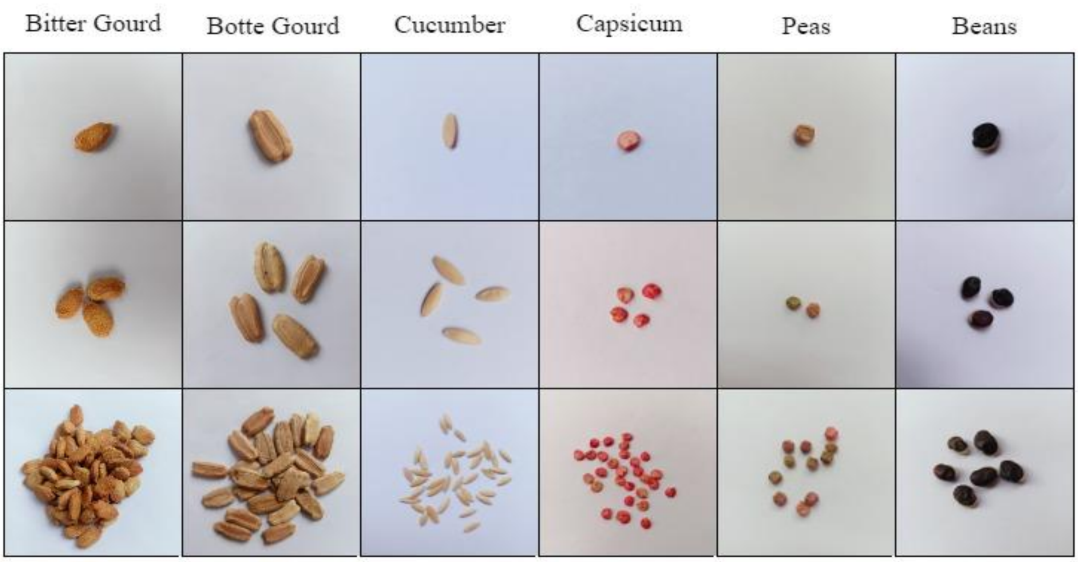

Figure 4.2.1. Images of seeds.

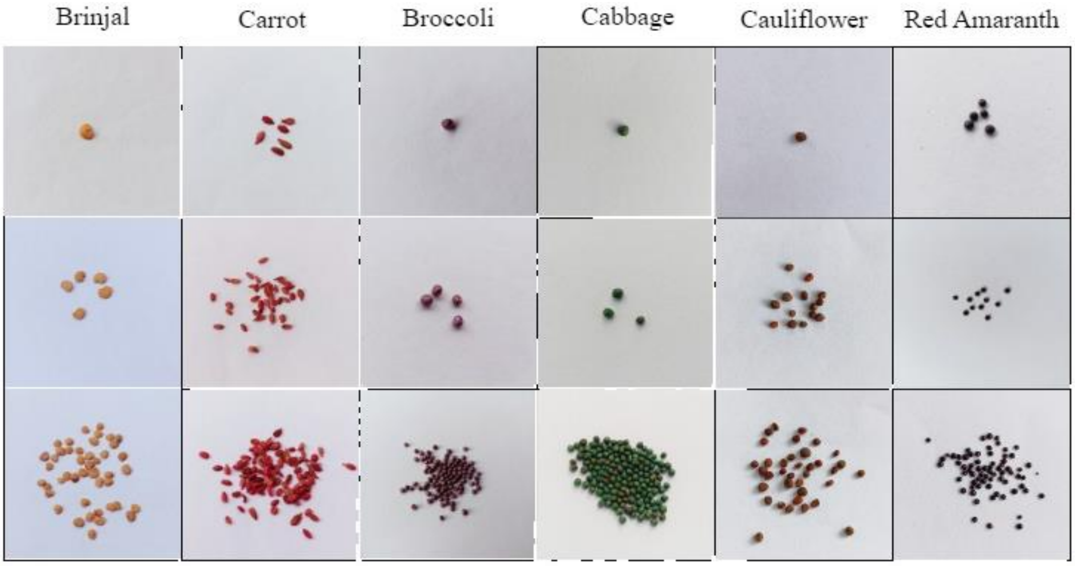

Figure 4.2.2. Images of seeds.

B. Dataset Cleaning and Pre-processing:

Initially, 6668 images were taken, representing twelve different kinds of vegetable seeds. Challenges such as random color changes and blurriness were seen throughout the image capturing process, which may compromise the quality of the dataset as a whole and the clarity of the images. To address this, images were manually removed, resulting in a standardized dataset of 6000 images, each containing 500 samples. The dataset was validated by an agronomist.

Due to the high-resolution photos being taken with mobile devices, there was a problem with computation time and efficiency. A rescaling approach was used to mitigate this. Using this method, the original pixel range of 0 to 255 was converted to a new range of 0 to 1 by normalizing the image pixel values by dividing each pixel's value by 255 [18]. Without changing the dimensions or shapes of the photos, this normalization ensured that all of the pixel values were within the same ranges.

The preprocessed images were then scaled to 300x300 pixels and 224x224 pixels in resolution. This resizing not only maintained the essential features required for accurate classification but also optimized the images for faster and more efficient processing by the deep learning models. These preprocessing procedures ensured high-quality input data and efficient computational performance, which was essential to making the dataset prepared for efficient model training and evaluation.

C. Dataset Splitting:

After preprocessing, the images were stored on Google Drive in folders according to their types, making them easily accessible for use in Google Colab. The dataset was divided into three parts, properly assessing the performance of the model: training, testing, and validation sets.

TABLE 4.2.2. DATASET SPLITTING.

Total	Training (80%)	Validation (10%)	Testing (10%)
6000	4800	600	600

In this study, 10% of the images were kept for testing, 10% were for validation, and the remaining 80% of the images were used to train the model. The model was fitted using the training set, tuned and improved using the validation set, and tested to assess the trained model's final performance [19].

4.3 Train Model

After preprocessing the dataset, transfer learning-based approach was utilized for seed variety identification, leveraging pre-trained models to extract relevant features from seed images. These features were then used to train a classification model, enabling the identification of seed varieties based on distinct characteristics such as color, size, and shape. Specifically, three deep learning models MobileNetV2, InceptionV3, and NASNetMobile were selected to analyze each model for complexity and computational efficiency, ultimately aiming to identify the most suitable one for seed classification.

The pre-trained models, loaded with ImageNet weights, had their final classification layers removed to retain only the convolutional base, which was used to extract high-level features from the seed images. These extracted features were then fed into a custom fully connected neural network trained to classify the seed varieties. The models were evaluated based on accuracy, complexity, and computational efficiency, with the best-performing model optimized and deployed for real-time seed variety identification. This transfer learning approach facilitated the application of knowledge from large datasets to improve performance on seed classification, enhancing both accuracy and efficiency.

Inception V3:

InceptionV3 was first developed as a module for Google's GoogLeNet, is designed to assist in image analysis. It is the third version of Google's Inception CNN series [20]. Allowing the building of deeper networks without significantly increasing the number of parameters was the main design objective of InceptionV3. This was performed using a number of innovative techniques.

One such method is the application of factorized convolutions, which reduces computational costs by replacing smaller convolutional filters for larger ones. This keeps complicated pattern recognition capabilities while minimizing the number of parameters and operations [20]. Additionally, the model uses significant normalization methods including label smoothing, dropout, and batch normalization to avoid overfitting and enhance generalization capabilities [21].

These advancements enable InceptionV3 to achieve outstanding performance and accuracy in image classification tasks while ensuring computational efficiency, which has resulted in its widespread adoption for a variety of computer vision applications, such as those in agriculture and healthcare.

NASNetMobile:

In 2018, Google Brain developed NASNetMobile, a convolutional neural network (CNN) that makes utilization of Neural Architecture Search (NAS). In order to optimize network designs, NAS is a gradient-based technique that makes use of reinforcement learning [22]. The search procedure provides NASNetMobile, which achieves a compromise between computational efficiency and performance, making it suitable for mobile device deployment.

A collection of operational blocks, including identity mappings, different convolutions, and pooling operations, are used to build the architecture [22]. The building components of the NASNet design are cells, which are created by combining these blocks. There are two different kinds of cells. Normal cells preserve the feature map's size, while reduction cells shrink its dimensions [22]. With 5,326,716 parameters, NASNetMobile achieves a parameter-efficient architecture compared to NASNetLarge's 88,949,818 parameters [23].

Proposed Model: MobileNetV2

MobileNet-V2 is a compact convolutional neural network (CNN) architecture of 53 layers [24]. This model was trained using over a million photos from the ImageNet collection [24]. It is made especially for embedded and mobile vision applications. The architecture of Mobilenetv2 contains a number of key elements that enhance its efficacy and efficiency in image classification. These features include linear bottlenecks, inverted residuals, bottleneck design, and depth wise separable convolution [25]. Each of these features is essential to keep the model's accuracy high while minimizing its computing complexity.

A. Depthwise Separable Convolution:

The depthwise separable convolution technique significantly reduces the computational cost of the convolutions by dividing the traditional convolution process into two more steps: depthwise convolution and pointwise convolution [25]. Depthwise convolution applies a single filter to each input channel individually, reducing the number of computations because each channel is processed separately. Subsequently, the pointwise convolution employs 1×1 convolutions to merge the outputs of the depthwise convolution in various channels, therefore efficiently integrating data from every channel [26]. The model becomes considerably more efficient while maintaining good performance as a result of this separation, which also significantly reduces the processing requires.

B. Linear Bottleneck:

Linear bottleneck is a deep neural network design method, primarily used in MobileNetV2, in which linear layers are integrated within convolutional blocks to retain information that could have been lost due to activation functions such as ReLU. ReLU can discard data by setting negative values to zero when it activates, but the network behaves linearly in the non-zero domains [26]. Because linear transformations can maintain all relevant data, linear bottlenecks ensure that important information is maintained during the dimensionality reduction process. Without losing important details, the network is still able to capture the necessary complexity by integrating the input manifold into a lower-dimensional subspace [25]. This approach ensures performance while lowering the size of the model and the amount of computing needed.

Designed for mobile and embedded devices, MobileNetV2's linear bottlenecks, improving efficiency and accuracy.

C. Inverted Residual:

Inverted residual blocks utilize a unique bottleneck construction to increase model accuracy. This structure started with an input, passes through a number of layers that are bottlenecks, and ends with an expansion step. The main novelty with inverted residuals, compared to conventional residual blocks, is the creation of bypasses between these bottlenecks.

Two main insights drive this design decision:

- The bottleneck layers capture the most significant information; hence the expansion layer is primarily an implementation detail for non-linear transformation.
- The training process is enhanced by shortcuts, which facilitate gradients' more efficient passing through several layers.

Prioritizing bottlenecks over expansion layers improves performance and uses less memory, which makes it suitable for devices with limited resources.

D. Model Architecture:

Though three models are employed in this study, MobileNetV2 is providing better results. Because, for image classification it is a lightweight and powerful model. The procedure in this model architecture begins with expanding the input into a higher-dimensional space from a compact low-dimensional representation. This expansion enables more detailed feature extraction. A quick depthwise convolution is then used. Depthwise convolution is computationally efficient because it conducts spatial convolution individually on each channel, resulting in a substantially lower number of parameters and computational cost than traditional convolutions.

After the depthwise convolution, a pointwise (1x1) convolution is applied. This stage converts the high-dimensional features back into a lower-dimensional space, efficiently compressing the data while retaining those essential features retrieved previously. Figure 4.3.1 Illustrate the architecture of MobileNetV2 [27].

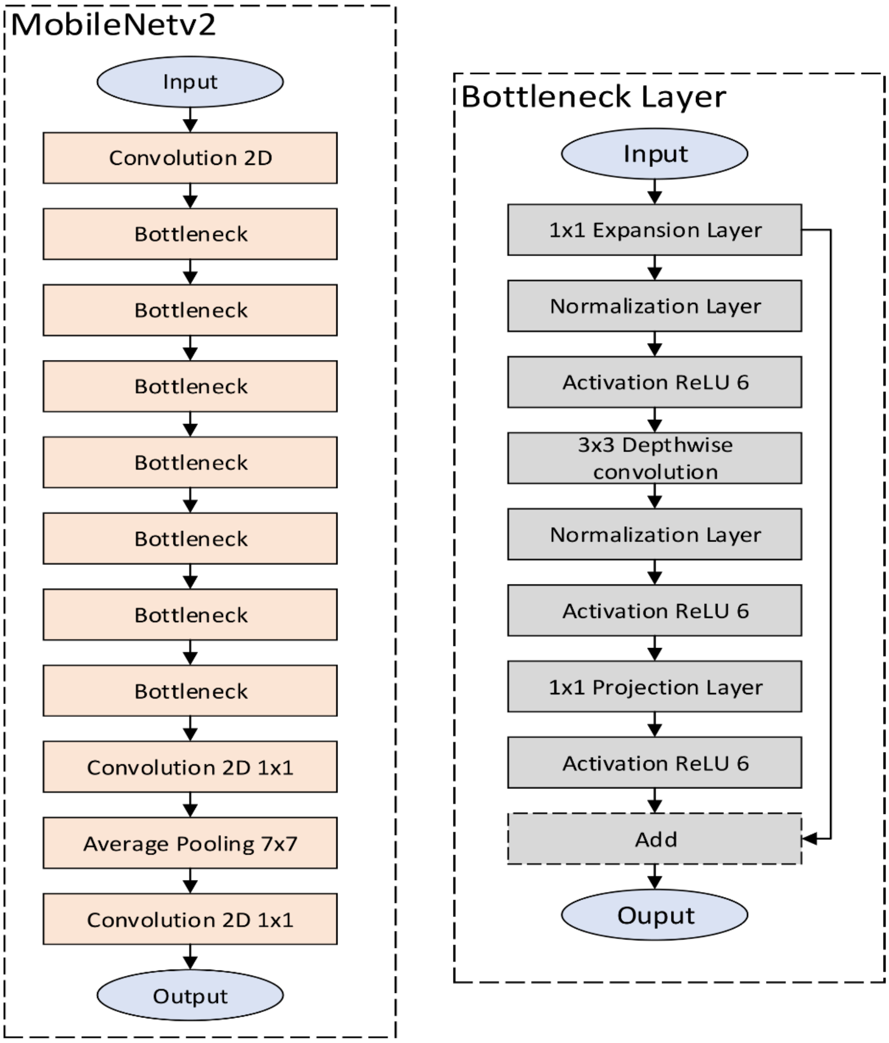

Figure 4.3.1: Architecture of MobileNetV2.

Compared to conventional techniques, this structural approach is intended to preserve information more successfully. By combining expansion, depthwise convolution, and compression, the model maintains a lightweight block structure that ensures efficiency and performance without compromising feature representation quality.

4.4 Model Evaluation

The evaluation of each model involves comprehensive statistical analysis using key performance metrics such as confusion matrix, accuracy, precision, recall, and F1 score.

Confusion Matrix: The confusion matrix provides a detailed breakdown of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), offering insights into the model's predictive capabilities.

TP: Correctly predicted positive cases.

TN: Accurately predicted negative cases.

FP: Incorrectly predicted positive cases.

FN: Incorrectly predicted negative cases.

Accuracy: Accuracy is the percentage of accurately predicted cases out of all instances [28]. It is the fundamental measure for evaluating model performance.

$$\text{Accuracy} = \frac{\mathrm{TP+TN}}{\mathrm{FP+FN+TP+TN}}$$

Precision: Precision is the proportion of true positive to overall positive predictions [28].

$$\text{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$$

Recall: Recall is the percentage of true positive based on actual positive instance [28].

$$\text{Recall} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$$

F1-Score: The F1 score is the harmonic mean of precision and recall [28].

$$\text{F1-Score} = 2 \times \frac{\mathrm{Precision \times Recall}}{\mathrm{Precision \times Recall}}$$

Accuracy and Loss Curves: Accuracy and loss curves represent the model's performance over epochs, presenting insight into training processes and potential overfitting or underfitting issues. The accuracy curve plots training and validation accuracy over epochs. The Loss Curve plots training and validation loss over epochs.

4.5 Summary

In conclusion, this research demonstrates the effectiveness of transfer learning in enhancing seed classification accuracy. MobileNetV2 emerges as the optimal model, offering a balance of computational efficiency and high classification accuracy. Statistical analyses including confusion matrices and accuracy curves provide insights into model behavior, supporting its potential application in improving agricultural practices through automated seed quality assessment. The study underscores the importance of leveraging advanced AI techniques to address critical challenges in agriculture, paving the way for enhanced crop productivity and management.

CHAPTER 5
Result and Analysis

5.1 Overview

The research focuses on optimizing vegetable seed classification using advanced deep learning techniques to aid agricultural practices in Bangladesh. This chapter details the experimental setup and evaluation criteria for three transfer learning models—MobileNetV2, InceptionV3, and NASNetMobile—based on their performance and computational efficiency. Experiments were conducted at resolutions of 224x224 and 300x300 pixels, over 20 and 100 epochs, to determine the optimal model. The study aimed to identify the model that balances accuracy and computational efficiency, providing a robust solution for practical applications in agriculture.

5.2 Experimental Result

During the experimental phase, the experiments are conducted based on image dimensions and epoch size while keeping other parameters constant, such as learning rate, dropout, and neuron number. Across all stages, three transfer learning models are implemented as previously described. Although all models perform adequately, this phase aims to analyze and determine which model is the most optimal and understand the reasons behind its performance.

TABLE 5.2.1. ACCURACY TABLE FOR EACH MODEL.

Model & Different Settings	**Number of Epoch 20**		**Number of Epoch 100**	
	224x224	**300x300**	**224x224**	**300x300**
MobileNetV2	98.67	99.17	99.33	99.50
Inception V3	98.33	98.50	98.67	98.50
NASNetMobile	98.50	98.67	98.17	99.17

In the experiments conducted at resolutions of 224x224 and 300x300, several models were evaluated over 20 epochs to determine their performance and computational efficiency.

For the 224x224 resolution, all models performed similarly across the epochs. Notably, MobileNetV2 performed exceptionally well, achieving its best result 98.50% at epoch 18, as shown in Figure 5.2.1. This indicates that MobileNetV2 offers a strong balance of accuracy and computational efficiency at this resolution.

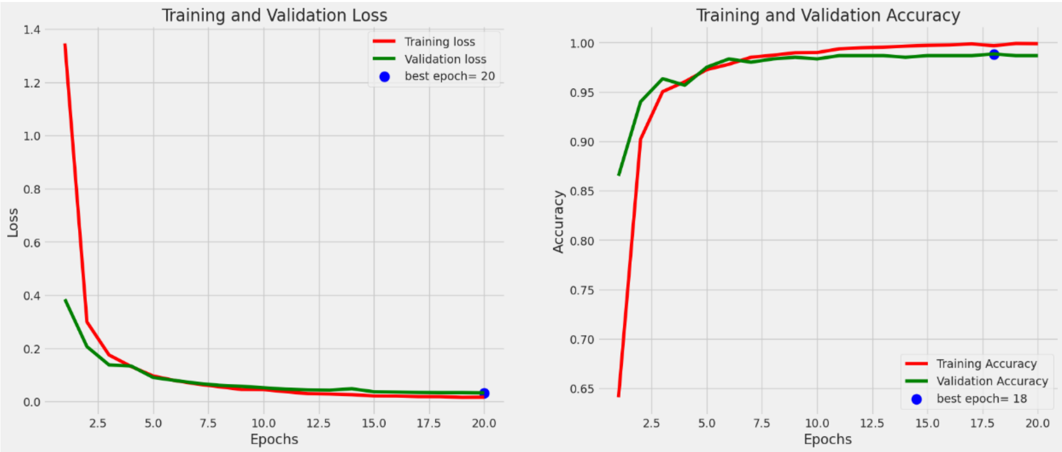

Figure 5.2.1: Training and validation accuracy and loss over the 20 epochs with 224x244 dimension.

For the 300x300 resolution, MobileNetV2 again demonstrated superior performance by reaching its peak accuracy of 99.33% at epoch 13, as shown in Table 5.2.2. This early peak is significant as it indicates less computational time and resources needed to achieve high performance. In comparison, Inception V3 and NASNetMobile achieved accuracies of 98.50% and 98.76% at epochs 16 and 20, respectively. Although these models achieved high accuracy, the extended training periods required to reach their best performance make them less efficient compared to MobileNetV2.

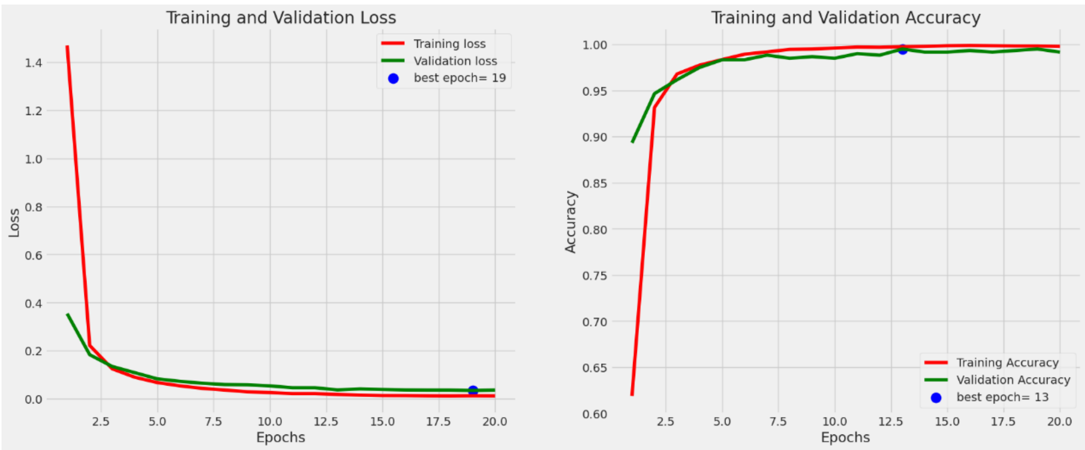

Figure 5.2.2: Training and validation accuracy and loss over the 20 epochs with 300x300 dimension.

Overall, while Inception V3 and NASNetMobile are capable of high accuracy, MobileNetV2 stands out for its superior balance of accuracy and computational efficiency. Achieving the highest accuracy of 99.33% by epoch 13 at a resolution of 300x300 makes MobileNetV2 particularly advantageous for practical applications requiring efficient and effective model deployment.

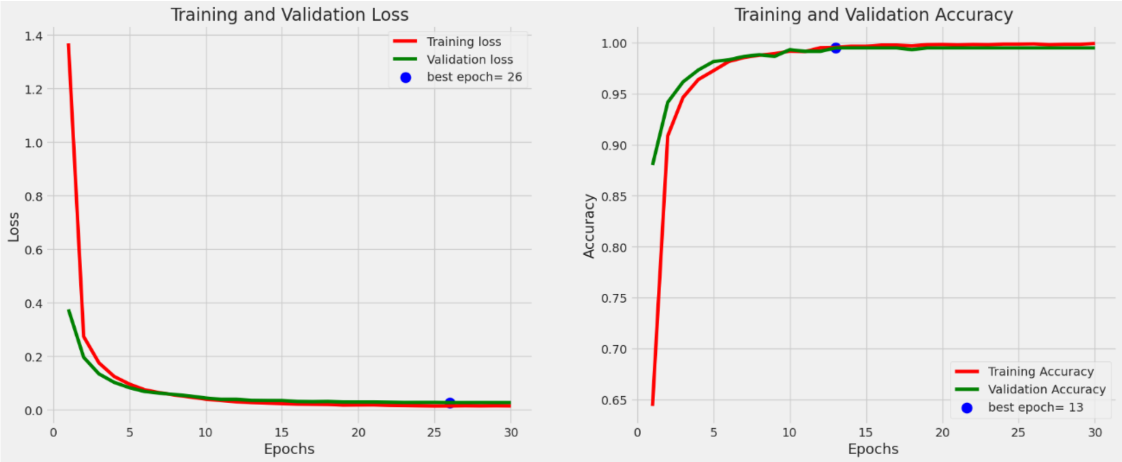

Figure 5.2.3: Training and validation accuracy and loss over the 100 epochs with 224x244 dimension.

In the experiments conducted at same resolution over the 100 epochs, all models were evaluated to observe changes in performance. MobileNetV2 consistently performed well at both resolutions. For the 224x224 resolution, MobileNetV2 achieved an accuracy of 99.50% at epoch 13, as shown in Figure 5.2.3. For the 300x300 resolution, MobileNetV2 reached an accuracy of 99.50% at epoch 10, as shown in Figure 5.2.4.

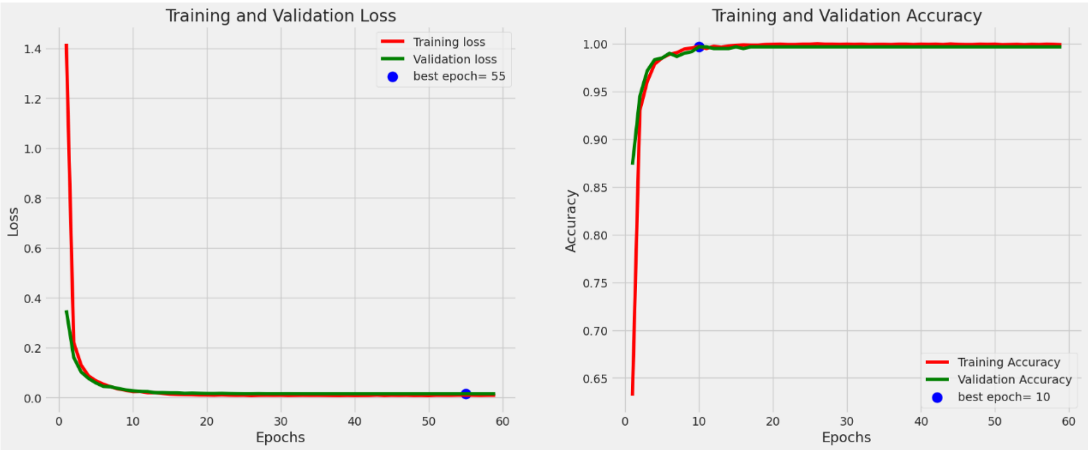

Figure 5.2.4: Training and validation accuracy and loss over the 100 epochs with 300x300 dimension.

NASNetMobile took significantly longer to achieve its best accuracy, making it computationally extensive. Inception V3 achieved an accuracy of 98.67% at epoch 26 for the 224x224 resolution and performed similarly for the 300x300 resolution across the 100 epochs, mirroring its performance over 20 epochs. These results highlight MobileNetV2's superior balance of accuracy and computational efficiency at both resolutions, reaching high performance much earlier than NASNetMobile and Inception V3.

TABLE 5.2.2. LOSS TABLE FOR EACH MODEL.

Model & Different Settings	**Number of Epoch 20**		**Number of Epoch 100**	
	224x224	**300x300**	**224x224**	**300x300**
MobileNetV2	3.17	3.42	2.74	1.57
Inception V3	5.50	3.80	5.61	3.71
NASNetMobile	7.47	4.15	7.91	4.17

The performance of the models depends not only on accuracy but also on loss, where lower loss indicates better performance. As shown in Table 5.2.2, MobileNetV2 experienced the least data loss during training compared to all other models. In contrast, NASNetMobile experienced the highest data loss, highlighting its inefficiency.

Training and validation accuracy are common criteria for determining a model's efficacy. During training, the model learns from the training dataset, and the accuracy achieved is referred to as training accuracy. However, training accuracy alone is insufficient for determining a model's predictive capabilities. Validation accuracy, which assesses the model's performance on a separate validation dataset, contributes to its capacity to generalize to new data. In addition, the model frequently computes a loss function while training. Training loss analyzes how well the model's predictions match the actual target values in the training dataset, where validation loss analyzes the model's performance on the validation dataset in a similar manner. Lower training and validation losses usually indicate better model performance.

Initially, training and validation accuracy improved, as shown in the figure 5.2.1, 5.2.2. 5.2.3, 5.2.4. After a few epochs, they began to converge, showing a balance of learning and generalization. As expected, the model consistently performed better on the training data. Furthermore, the training and validation losses gradually decreased during training, which is beneficial. The training loss consistently lower than the validation loss, showing high adaptability. The above figures demonstrate that the training and validation data are properly separated, with no overfitting visible.

5.3 Comparative Analysis

From the previous discussion, it is evident that MobileNetV2 performs well in all stages. Several factors contribute to its superior performance because of MobileNetV2 is streamlined, quick, and energy-efficient, making it highly effective. In this section, describing a comparison of MobileNetV2's performance based on epoch and dimension settings.

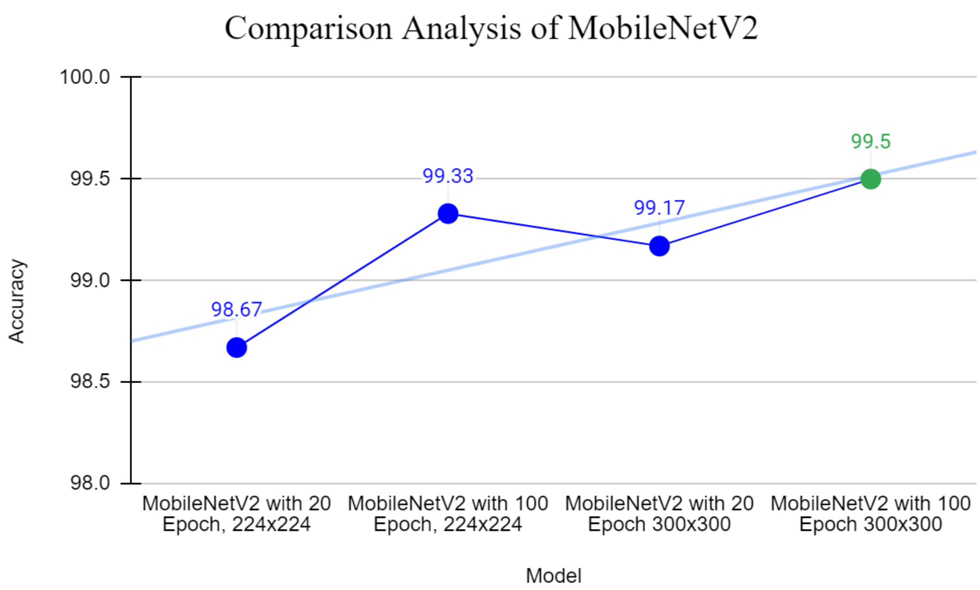

Figure 5.3.1: Comparison analysis of MobileNetV2 based on different epoch and dimension.

Increasing the image resolution to 300x300 allows the model better understanding of images because it contains more detailed information than a lower resolution, such as 224x224. This additional information can be valuable in uses involving delicate analysis or the identification of hidden features in images. When the model trains on better-quality images, it is able to extract more detailed patterns and characteristics with each epoch. Essentially, each training iteration receives more input knowledge, allowing the model to learn more discriminative features and increase overall performance.

Moreover, increasing the number of epochs improves the model's learning capability. With more epochs, the model has more possibilities to fine-tune its interpretation of the data, improve its prediction accuracy, and generalize better to new instances. Increasing resolution and epochs improves model accuracy but extends training time due to increased processing requirements. Finding the appropriate balance is crucial for successful analysis of images. Figure 5.3.1 illustrates that MobileNetV2 achieves its best performance with higher image dimensions (300x300) over 100 epochs, achieving an accuracy of 99.50%. The corresponding loss, as shown in Table 5.2.2, is minimized at 1.57%.

Classification report is a commonly used tool for evaluating the performance of a model of classification. It summarizes key indicators for assessing a model's ability to accurately classify data, such as precision, recall, F1-score, and support that are previously described. High precision and recall scores are often preferred, showing that the model can accurately recognize instances of each category. Table 5.3.1 shows a high level of accuracy and efficacy in classifying the various categories, as mentioned in the classification report.

TABLE 5.3.1. CLASSIFICATION REPORT.

Category	**Precision**	**Recall**	**F1-Score**	**Support**
Beans	1.00	1.00	1.00	54
Bitter Gourd	1.00	1.00	1.00	38
Bottle Gourd	1.00	1.00	1.00	59
Brinjal	1.00	1.00	1.00	53
Broccoli	0.98	0.98	0.98	41
Cabbage	1.00	1.00	1.00	53
Capsicum	1.00	1.00	1.00	49
Carrot	1.00	1.00	1.00	54
Cauliflower	1.00	0.98	0.99	49
Cucumber	1.00	1.00	1.00	51
Peas	1.00	1.00	1.00	45
Red Amaranth	0.98	1.00	0.99	54

The confusion matrix is a visual representation of a classification model's performance, comparing actual and predicted classifications for every class. Figure 5.3.2 shows errors in classifying broccoli, cauliflower, and red amaranth seeds, whereas the other categories are correctly identified.

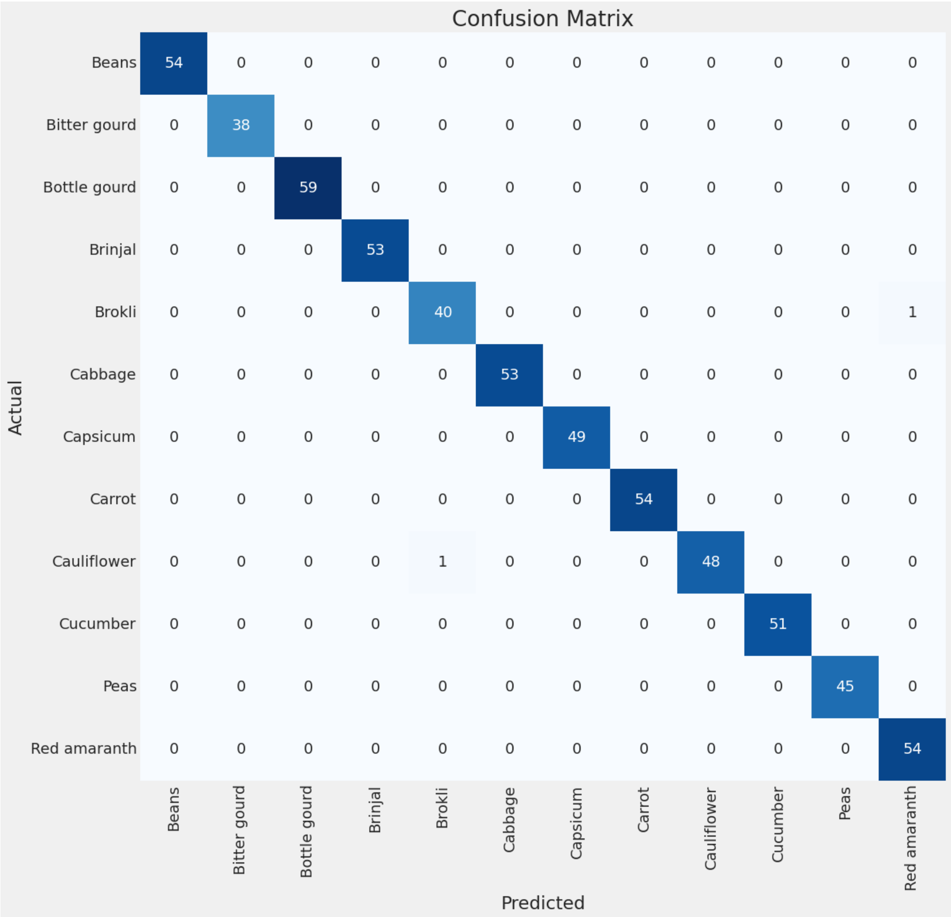

Figure 5.3.2: Confusion Matrix.

5.4 Summary

MobileNetV2 emerged as the most efficient model for vegetable seed classification, outperforming InceptionV3 and NASNetMobile. At a resolution of 300x300 pixels over 100 epochs, MobileNetV2 achieved the highest accuracy of 99.50% and the lowest loss of 1.57%. The comparative analysis highlighted MobileNetV2's superior performance due to its streamlined architecture and computational efficiency. The model demonstrated a strong balance of accuracy and processing time, making it ideal for real-time applications in agriculture. The classification report and confusion matrix further validated MobileNetV2's effectiveness, showcasing high precision, recall, and F1 scores across all seed categories.

CHAPTER 6

Impact on Society, Environment and Sustainability

6.1 Impact on Life

The implementation of image processing-based transfer learning for vegetable seed classification in Bangladesh has significantly improved the lives of farmers and communities. By providing accurate and reliable seed identification tools, the technology enhances agricultural efficiency, leading to increased crop yields and improved food security. Farmers benefit from higher incomes due to better seed quality, which helps alleviate poverty and stabilize rural economies. Moreover, democratizing agricultural knowledge through accessible technology empowers diverse demographics, including younger generations and urban dwellers interested in farming, fostering a more inclusive and knowledgeable farming community.

6.2 Impact on Society, Environment

The implementation of an image processing-based transfer learning approach for the classification of vegetable seeds in Bangladesh has a number of societal implications. Primarily, it improves agricultural efficiency by giving farmers with an accurate and trustworthy technique of seed identification. This accuracy guarantees that farmers cultivate proper seeds, resulting in increased crop yields and food security. Improved seed production and quality can boost farmers' incomes, helping reduce poverty and improve economic stability in rural communities. Furthermore, this technology has the potential to democratize agricultural knowledge, making it available not only to seasoned farmers, but also to younger people and city dwellers interested in urban farming. By building a more aware farming community, the project encourages sustainable agricultural practices and increases overall agricultural productivity.

Overall, implementing this advanced technology in Bangladesh's agriculture sector represents a transformative opportunity. It not only increases efficiency and productivity, but it also has a significant impact on economic stability, poverty reduction, sustainable practices, and the well-being of agricultural communities. This project provides the foundation for a more inclusive, informed, and sustainable agricultural future by making innovative agricultural technology available to more people.

The project has a significant positive environmental impact. The project promotes the utilization of high-quality seeds through correct classification, resulting in more efficient use of land and resources in agriculture. High-quality seeds often produce healthier plants that are more resistant to pests and diseases, decreasing the need for chemical inputs like fertilizer and insecticides. This reduction not only lowers farmers' production expenses, but also reduces the environmental impact of agricultural activities. By optimizing seed selection, the research helps to encourage sustainable farming techniques that save natural resources and maintain environmental balance. Furthermore, accurate seed classification is crucial for maintaining biodiversity and environmental stability.

In conclusion, the use of advanced seed classification technology not only increases agricultural productivity and economic consequences, but it also plays an important role in supporting environmental sustainability. By encouraging the use of high-quality seeds and minimizing chemical inputs, the project helps to create a healthier environment, preserve biodiversity, and encourage adaptable agricultural systems that can adapt to changing environmental conditions.

6.3 Ethical Aspects

The study emphasizes ethical considerations. The approach prioritizes precision and reliability in seed classification to help the farming community. Informed consent is required in order to ensure that data gathering respects participants' rights and privacy. Transparency is essential, with full documentation allowing criticism and reproducibility, ensuring that the benefits of the research are broadly available. The project also takes consideration of the socioeconomic background, with the goal of improving food security and economic stability while minimizing current disparities.

This commitment to sustainable development contributes to the preservation of biodiversity and natural resources while also increasing agricultural output. Overall, the ethical considerations of this research ensure that developments in agricultural technology help both social well-being and environmental sustainability.

6.4 Sustainability Plan

To assure the project's sustainability, a multifaceted approach is required. For example, cooperation with local agricultural organizations and government agencies can help the technology become more widely adopted and integrated into existing agricultural practices. Farmers will benefit from ongoing training and support programs, ensuring that they can use technology efficiently. Furthermore, creating a business model with inexpensive pricing and potential subsidies for low-income farmers can improve accessibility. Collaboration with academic and research institutes can improve the system through continuous research and innovation. Finally, supporting environmental sustainability through responsible resource usage and minimizing the technology's carbon footprint will assure its long-term viability.

6.5 Summary

This study has brought transformative impacts across various aspects of agriculture in Bangladesh. It significantly enhances agricultural efficiency, leading to increased crop yields and improved food security. By democratizing agricultural knowledge, the technology empowers farmers and promotes sustainable practices, thereby contributing to economic stability and poverty reduction in rural communities. Moreover, the project has a positive environmental impact by promoting the use of high-quality seeds and reducing chemical inputs, thereby supporting biodiversity and environmental sustainability. Ethical considerations ensure transparency and respect for participants' rights, while a robust sustainability plan ensures long-term viability through collaboration, innovation, and responsible resource management. Overall, the project lays the foundation for a more inclusive, informed, and sustainable agricultural future in Bangladesh.

CHAPTER 7
Conclusion and Future Work

7.1 Conclusion

The study aims to improve the classification of vegetable seeds in Bangladesh by employing an image processing-based transfer learning method. The agricultural industry plays an important role to the country's economic development, and seed quality influences crop productivity and quality. The study acknowledged the difficulties faced by not only Bangladeshi farmers but also rural dwellers due to the complexity of manually identifying seed types. To solve these challenges, an automated method that uses transfer learning was proposed to improve the efficiency and accuracy of seed classification.

Three transfer learning models—MobileNetV2, Inception V3, and NASNetMobile—are evaluated for their capability to classify 12 types of vegetable seeds. High-resolution seed images have been obtained, preprocessed, and utilized to train the models. The study focused on enhancing the models by altering image resolutions and epoch parameters to determine the suitable configuration for accurate classification.

The experimental results demonstrated that MobileNetV2 performed better than the other models in terms of accuracy and computing efficiency. It achieved an outstanding precision of 99.50% at 300x300 resolution over the 100 epoch, indicating its suitability for practical applications. The study also stressed the significance of effective data preparation, such as rescaling and normalization, in ensuring high-quality input data for model training.

7.2 Future Suggested Work

The outcomes of the research highlight the potential of AI and transfer learning to revolutionize farming. Future research could look into the scalability and flexibility of the established model across different geographical regions and crop varieties, hence increasing its value and robustness in real-world agricultural contexts. Furthermore, efforts should be made to integrate this technology into user-friendly platforms such as smartphone applications, enabling accessibility and ease of use for farmers of diverse technological ability.

Expanding the dataset to include a wider variety of seed varieties, as well as more complicated background variables, might help to improve the model's adaptability and efficiency in real-world scenarios. Collaborative research including agronomists, data scientists, and local farmers would be crucial in improving the system and demonstrating its efficacy in various agricultural contexts.

Addressing these constraints in future studies, such as expanding the dataset to include more diverse backgrounds and real-time seed variety classification from larger sets, would be critical for advancing AI models in agricultural applications. Furthermore, alternate approaches, such as fine-tuning transfer learning models or creating custom CNN models, may improve in seed classification tasks. These advancements have the potential to significantly transform farming activities, benefiting both farmers and the entire agricultural sector.

7.3 Limitations

This research has several specific limitations. Firstly, the dataset of 6000 images, while covering twelve seed categories, may not fully represent the variability found in real-world conditions, potentially affecting the model's ability to generalize. Moreover, the study's reliance on three specific pre-trained models (MobileNetV2, Inception V3, and NASNetMobile) limits exploration of other potentially effective architectures that could offer different trade-offs in accuracy and computational efficiency. Additionally, the restriction to capturing seed images on a white background overlooks the variability introduced by real-world environmental conditions, potentially limiting the model's adaptability. Finally, the study's reliance on pre-trained models precludes the exploration of custom-built architectures tailored to specific nuances of seed classification. Addressing these limitations could involve expanding the dataset to include more diverse environmental conditions and seed variations, exploring alternative model architectures, and optimizing models for deployment on resource-constrained devices.

References:

[1] Husain, K. A. (2016, May 11). Role of quality seed in agriculture. The Daily Star. https://www.thedailystar.net/round-tables/role-quality-seed-agriculture-1222168

[2] Badikheti. (2023, March 29). What is the importance of seeds in agriculture?. badikheti. https://www.badikheti.com/resources/importance-of-seeds/

[3] Ataur, & An Overview of Automotive industry in Bangladesh | Expert tips. (2021, May 29). Major Crops of Bangladesh. BD Economy. https://oikosmist.com/crops-of-bangladesh/

[4] Fao.org. Bangladesh | FAO Regional Office for Asia and the Pacific | Food and Agriculture Organization of the United Nations. (n.d.). https://www.fao.org/asiapacific/perspectives/agricultural-statistics/global-strategy/results-in-the-region/bangladesh/en/

[5] Zhu, S., Zhang, J., Chao, M., Xu, X., Song, P., Zhang, J., & Huang, Z. (2019). A rapid and highly efficient method for the identification of soybean seed varieties: hyperspectral images combined with transfer learning. Molecules, 25(1), 152.

[6] Koklu, M., & Ozkan, I. A. (2020). Multiclass classification of dry beans using computer vision and machine learning techniques. *Computers and Electronics in Agriculture*, *174*, 105507.

[7] Rahman, M. M., & Barua, P. (2021, December). A CNN Model-based ensemble approach for Fruit identification using seed. In 2021 5th International Conference on Electrical Information and Communication Technology (EICT) (pp. 1-6). IEEE.

[8] Singh, A. (2023, November 22). A comprehensive guide to ensemble learning (with python codes). Analytics Vidhya. https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/

[9] Khatri, A., Agrawal, S., & Chatterjee, J. M. (2022). Wheat seed classification: utilizing ensemble machine learning approach. Scientific Programming, 2022.

[10] Kiratiratanapruk, K., Temniranrat, P., Sinthupinyo, W., Prempree, P., Chaitavon, K., Porntheeraphat, S., & Prasertsak, A. (2020). Development of paddy rice seed classification process using machine learning techniques for automatic grading machine. Journal of Sensors, 2020.

[11] Tuğrul, B. (2022). Classification of five different rice seeds grown in Turkey with deep learning methods. Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering, 64(1), 40-50.

[12] Keya, M., Majumdar, B., & Islam, M. S. (2020, July). A robust deep learning segmentation and identification approach of different bangladeshi plant seeds using CNN. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.

[13] Gulzar, Y., Hamid, Y., Soomro, A. B., Alwan, A. A., & Journaux, L. (2020). A convolution neural network-based seed classification system. Symmetry, 12(12), 2018.

[14] Seethalakshmi, A. V., & Hemachitra, H. S. (2020). Complex type seed variety identification and recognition using optimized image processing techniques. ACCENTS Transactions on Image Processing and Computer Vision, 6(19), 23.

[15] Vijayakanthan, G., Kokul, T., Pakeerathai, S., & Pinidiyaarachchi, U. A. J. (2021, August). Classification of vegetable plant pests using deep transfer learning. In 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS) (pp. 167-172). IEEE.

[16] Jaithavil, D., Triamlumlerd, S., & Pracha, M. (2022, March). Paddy seed variety classification using transfer learning based on deep learning. In 2022 International Electrical Engineering Congress (iEECON) (pp. 1-4). IEEE.

[17] Agarwal, D., & Bachan, P. (2023). Machine learning approach for the classification of wheat grains. Smart Agricultural Technology, 3, 100136.

[18] Team, K. (no date) Keras Documentation: Rescaling Layer, Keras. Available at: https://keras.io/api/layers/preprocessing_layers/image_preprocessing/rescaling/ (Accessed: 17 May 2024).

[19] Training, validation, test split for Machine Learning Datasets (no date) Encord. Available at: https://encord.com/blog/train-val-test-split/ (Accessed: 17 May 2024).

[20] *INCEPTIONV3* (2024) *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Inceptionv3 (Accessed: 25 May 2024).

[21] Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). pmlr.

[22] Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8697-8710).

[23] Team, K. (no date) *Keras Documentation: Keras applications*, *Keras*. Available at: https://keras.io/api/applications/ (Accessed: 26 May 2024).

[24] *ImagePretrainedNetwork* (no date) *(Not recommended) MobileNet-v2 convolutional neural network - MATLAB*. Available at: https://www.mathworks.com/help/deeplearning/ref/mobilenetv2.html (Accessed: 26 May 2024).

[25] Sharma, N. (2024) *What is mobilenetv2? features, architecture, application and more*, *Analytics Vidhya*. Available at: https://www.analyticsvidhya.com/blog/2023/12/what-is-mobilenetv2/ (Accessed: 27 May 2024).

[26] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).

[27] Tragoudaras, A. (2022) Design space exploration of a sparse mobilenetv2 using high-level synthesis and sparse matrix techniques on fpgas, MDPI. Available at: https://www.mdpi.com/1424-8220/22/12/4318 (Accessed: 27 May 2024).

[28] B, H.N. (2020) *Confusion matrix, accuracy, precision, recall, F1 score*, *Medium*. Available at: https://medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd (Accessed: 27 May 2024).

APPENDIX A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title:
IMAGE PROCESSING-BASED TRANSFER LEARNING APPROACH FOR CLASSIFICATION OF VEGETABLE SEEDS OF BANGLADESH.

Student ID: 203-15-14510

CO Description for FYDP

CO	CO Descriptions	PO
Phase -I		
CO1	Integrate recently gained and previously acquired knowledge to identify a **vegetable seeds classification** problem for the Final Year Design Project (FYDP).	**PO1**
CO2	Analyze different aspects of the goals in designing a solution for this FYDP	**PO2**
CO3	Explore diverse problem domains through a literature review, delineate the issues, and establish these goals for the FYDP	**PO4**
CO4	Perform economic evaluation and cost estimation and employ suitable project management procedures throughout the development life cycle of the FYDP	**PO11**
Phase -II		
CO5	Design and develop technical solutions and system components or processes that meet specified requirements, ensuring compliance with public health and safety standards, as well as considering cultural, socioeconomic, and environmental factors in this FYDP	**PO3**
CO6	Choose and apply appropriate methodologies, resources, and contemporary engineering and IT technologies to address complex engineering processes, encompassing prediction and modeling, while adhering to relevant constraints in this FYDP	**PO5**
CO7	Analyze societal, health, safety, legal, and cultural considerations, along with associated responsibilities, in the context of professional engineering practice and the resolution of this problem, employing logical reasoning guided by contextual understanding.	**PO6**
CO8	Comprehend and evaluate the enduring sustainability and impact of professional engineering endeavors in addressing intricate engineering challenges within social and environmental frameworks.	**PO7**
CO9	Implement ethical principles and adhere to professional standards and norms in this FYDP	**PO8**
CO10	Capable of operating proficiently both individually and as a team member or leader across diverse teams and interdisciplinary settings in this FYDP.	**PO9**

CO11	Proficiently communicate with the engineering community and broader society regarding complex engineering endeavors, including the ability to comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP.	**PO10**
CO12	Acknowledge the importance of self-directed and life-long learning within the evolving landscape of technology, and possess the readiness and capability to engage in lifelong learning endeavors.	**PO12**

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (with Knowledge Profile)	References
1.	**EP1: Depth of Knowledge required**	Yes	CO1, CO2, CO3, CO5, CO6, CO7 and CO8	For this project, I require theoretical knowledge about agricultural science and machine learning which covers **theory based natural science** (**K1).**	**Page no:** [1-2] **Section:** [1.2]
				This project demonstrates **fundamental engineering (K3)** principles by employing data collection and preprocessing. The project demonstrates **specialist knowledge (K4)** by conducting deep learning-based model like MobileNetV2, Inception V3 and NasNetMobile to enhance seed classification accuracy.	**Page no:** [13-16] **Section:** [4.2] **Page no:** [11-12] **Section:** [3.3]
				To complete this project, I need to understand the requirements of components and create a proper workflow that covers **engineering practice & design (K5).** This project addresses **engineering practice & technology (K6)** by employing deep learning models.	**Page no:** [10-11] **Section:** [3.2] **Page no:** [11-12] **Section:** [3.3]

				This project ensures to **K8 (Research Literature)** by reviewing existing studies with similar objectives.	**Page no:** [5-8] **Section:** [2.2, 2.3]
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	This project addresses **EP-2** by facing some challenges to making custom dataset. I need to conduct online research to find reliable sources of high-quality seeds. Then, consult experts from nurseries for insights.	**Page no:** [9] **Section:** [2.4]
3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	This project addresses **EP-3** by thoroughly comparing experimental outcomes, identifying Transfer Learning as the optimal solution for enhancing vegetable seed classification among various potential approaches.	**Page no:** [22-25] **Section:** [5.2]
4.	**EP4: Familiarity of Issues**	Yes	CO8	To work on seed image classification, I need to study agricultural aspects, despite not being familiar with them, and figure out how to classify seeds based on color, shapes, and size which indicates **EP-4**.	**Page no:** [1-2] **Section:** [1.2]
5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A

6.	**EP6: Extends of stakeholders involved and conflicting requirements**	No	CO8	N/A	N/A
7.	**EP7: Interdependence**	Yes	CO5	This project addresses **EP-7** by integrating data collection, statistical analysis, and methodology to provide a comprehensive solution for vegetable seed classification.	**Page no:** [10-11, 13-16] **Section:** [3.2-4.2]

Addressing CO11 with Complex Engineering Activities (EA):

SN	EA Definition	Attainment	CO	Justification	References
1.	**EA1: Range of resources**	Yes	CO11	This project leverages high-performance computing infrastructure, GPUs, deep learning frameworks, annotated datasets, and ethical guidelines to conduct systematic research and advance vegetable seed classification using transfer learning.	**Page no:** [11-12] **Section:** [3.3]
2.	**EA2: Level of interaction**	No		N/A	N/A
3.	**EA3: Innovation**	No		N/A	N/A
4.	**EA4: Consequences for society and the environment**	Yes		This project contributes to society by improving agricultural efficiency, promoting sustainable practices through accurate and reliable vegetable seed classification.	**Page no:** [29-31] **Section:** [6.1-6.4]
5.	**EA-5: Familiarity**	Yes		This project expands existing research by exploring transfer learning models for seed classification. It is supported by initial findings and a detailed comparative analysis, introducing innovative perspectives to the field.	**Page no:** [5-8] **Section:** [2.1-2.3]

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	This project addresses **CO4** by integrating effective project management and financial oversight. I prepared a budget to evaluate and estimate the cost required for this project.	**Page no:** [12] **Section:** [3.4]
2	CO9	Yes	This project prioritizes ethical standards in seed classification, ensuring precision, participant privacy, transparency, and socioeconomic equity. It aims to enhance agricultural productivity while promoting social welfare and environmental sustainability which comply with **CO9**.	**Page no:** [30] **Section:** [6.3]
3	CO10	No	N/A	N/A
4	CO12	Yes	The project's dedication to continuous learning (CO12) and adaptation in an evolving technological context is demonstrated through its meticulous data gathering, statistical analysis, methodology refinement, and comprehensive experimental results. This underscores its commitment to remaining up-to-date and improving strategies to effectively address current challenges in seed classification.	**Page no:** [10-11, 13-21, 22-28] **Section:** [3.2, 4.2, 4.3, 4.4, 5.2, 5.3]

APPENDIX B

[1] Shahriar, M. A., Rashmi, M. K., & Ahamed, M. S. (2023, December). Transfer Learning Approach for Classification of Popular Mango Species of Bangladesh. In *2023 26th International Conference on Computer and Information Technology (ICCIT)* (pp. 1-6). IEEE.

IMAGE PROCESSING-BASED TRANSFER LEARNING APPROACH FOR CLASSIFICATION OF VEGETABLE SEEDS OF BANGLADESH

ORIGINALITY REPORT

24%	21%	12%	17%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	6%
2	Submitted to Daffodil International University Student Paper	4%
3	iieta.org Internet Source	1%
4	www.mdpi.com Internet Source	1%
5	ebin.pub Internet Source	1%
6	dergipark.org.tr Internet Source	1%
7	Submitted to AlHussein Technical University Student Paper	1%
8	export.arxiv.org Internet Source	<1%
9	Submitted to Washington University in St. Louis	<1%