Design and Development of Seed Germination Classification System using AI and IoT

Submitted in partial fulfillment of the requirements

for the award of the degree of

DOCTOR OF PHILOSOPHY

Submitted by

D. Ramesh Reddy

(Roll No. 701944)

Under the supervision of

Prof. Prakash Kodali

Assistant Professor (Gr-I)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY WARANGAL
TELANGANA - 506004, INDIA

May 2024

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY WARANGAL
TELANGANA - 506004, INDIA

THESIS APPROVAL FOR Ph.D.

This is to certify that the thesis entitled, Design and Development of Seed Germination Classification System using AI and IoT, submitted by Mr. D. Ramesh Reddy [Roll No. 701944] is approved for the degree of DOCTOR OF PHILOSOPHY at the National Institute of Technology Warangal,Telanagana, India

Examiner

Research Supervisor
Prof. Prakash Kodali
Assistant Professor (Gr-I)
Dept. of Electronics and Communication Enggineering
NIT Warangal
India

Chairman
Prof. D. Vakula
Professor & Head
Dept. of Electronics and Communication Enggineering
NIT Warangal
India

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY WARANGAL
TELANGANA - 506004, INDIA

CERTIFICATE

This is to certify that the thesis entitled, Design and Development of Seed Germination Classification System using AI and IoT , submitted in partial fulfillment of the requirement for the award of the degree of DOCTOR OF PHILOSOPHY to the National Institute of Technology Warangal, is a bonafide research work done by Mr. D. Ramesh Reddy [Roll No. 701944] under my supervision. The contents of the thesis have not been submitted elsewhere for the award of any degree.

Research Supervisor
Prof. Prakash Kodali (Gr-I)
Department of Electronics and Communication Engineering
NIT Warangal
India

Place: NIT Warangal

Date: $1^{st} May, 2024$

DECLARATION

This is to certify that the work presented in the thesis entitled "*Design and Development of Seed Germination Classification System using AI and IoT* " is a bonafide work done by me under the supervision of Prof. Prakash Kodali was not submitted elsewhere for the award of any degree.

I declare that this written submission represents my ideas in my own words and where other's ideas or words have been included, I have adequately cited and referenced the sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented fabricated or falsified any idea/date/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

D. Ramesh Reddy
(Roll No. 701944)
Date: $1^{st} May, 2024$

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and appreciation to my supervisor, **Prof. Prakash Kodali** for his invaluable guidance throughout the completion of this work. Their continuous support, timely feedback, and constructive discussions have played a pivotal role in helping me achieve my objectives. I am grateful for the ample time they dedicated to reviewing my work and providing insightful suggestions for improvement. Their mentorship not only shaped me as a researcher but also as an individual. I have been inspired by their words, actions, and values, which have demonstrated the qualities of a great teacher and a compassionate human being. Their unwavering dedication and commitment to excellence have left a profound impact on me. I aspire to embody these remarkable qualities throughout my life.

I would like to express my heartfelt gratitude to all the members of my Doctoral Scrutiny Committee (DSC), namely Prof. D. Vakula, Prof. N. Bheema Rao, Prof. T.V.K. Hanumantha Rao, and Prof. Abdul Azeem P. Their valuable comments and suggestions during the oral presentations have enriched my research work. I am truly fortunate to have had the opportunity to attend lectures by esteemed professors such as Prof. T. Kishore Kumar, Prof. Ravi Kumar J, Prof. Lakshmi, and Prof. Muralidhar. Their knowledge and expertise have been instrumental in broadening my understanding of the field.

I am immensely thankful to Prof. Anjaneyulu, Prof. Sree Hari Patri, and Prof. Vakula D Heads of Dept. of ECE and chairman of DSC, during my tenure for providing adequate facilities in the department for carrying out the oral presentations. I wish to express my thanks to all the esteemed faculty members of the Department of Electronics and Communication Engineering at NIT Warangal. I would also like to extend my heartfelt gratitude to Prof. N.V. Ramana Rao and Prof. Bidyadhar Subudhi, the Director of NIT Warangal, for their unwavering support and encouragement throughout my research endeavor. I am truly fortunate to have been part of such a remarkable institution and to have received support from all these individuals who have contributed to my academic and personal development.

I would like to express my heartfelt gratitude to my senior Ravi Sankar Puppala as well as my co-scholars, Venkata Narayana, Ragini, Vijayasree, Lakshkoti and Sowjanya for their invaluable support and selfless cooperation throughout my Ph.D. journey. Their presence and assistance have been truly remarkable, and I am deeply thankful for their unwavering support and readiness to help me at any time. I would like to express a special thanks to Prof. Ramalingaswamy Cheruku for his unconditional support, and constant presence in both the joyful and challenging moments of my personal life and research work. I am forever grateful to have such amazing individuals by my side. Without the support and camaraderie of these remarkable individuals, my Ph.D. journey would have been much more challenging. I am truly blessed to have such incredible friends in my life, and I will cherish their friendship and support forever. I am immensely grateful to my wife G.Sneha, beloved son D. Jatin Reddy, my mother D. Radha and my father D. Raghu Rami Reddy, for their unwavering love, support, and prayers throughout my journey. Their constant encouragement and support have been my pillars of strength, providing me with the confidence and motivation to overcome challenges and pursue excellence.

D.Ramesh Reddy

Dedicated to

Family, Friends & Teachers

ABSTRACT

In the field of agriculture, seeds play a crucial role since they enable the domestication of a wide variety of plants and the creation of technical developments in the field of biotechnology. The quality of seeds, including their physical, genetic and physiological attributes, is crucial for their expression and impact on agricultural processes. High physiological quality seeds have influenced the development of major crops and have been essential in meeting the demands of a growing world population. Germination testing is crucial for evaluating seed quality, but it may be challenging. Precise environmental conditions are necessary for the procedure, and they must be carefully regulated to guarantee precision. The germination test findings might be influenced by temperature and humidity. Existing growth chambers are used to maintain a uniform temperature and humidity level, facilitating the germination of seeds. The prolonged time frame for germination testing, which may span from several days to weeks, is an additional concern, possibly causing delays in making germination testing. The manual examination in seed testing laboratories, the limited number of quality testing facilities makes the time limitation even more severe, creating a bottleneck that makes it impossible to scale up and fill the demand for rapid testing and agricultural decision making.

The proposed system is an automated growth chamber using Artificial Intelligence (AI) and Internet of Things (IoT) that will precisely regulate temperature and humidity to ensure optimal seed germination. The purpose of this method is to enhance the accuracy of predicting seed germination, hence facilitating the evaluation of seed quality, thereby enhancing agricultural production. Proposed system also assists in collection of large datasets. These datasets are essential for developing and improving AI models that predict seed germination. A novel two- stage network is developed that uses multiple Convolutional Neural Networks (CNN) to automate seed recognition and evaluation of germination condition evaluation. The Detectron2 framework is employed in the first stage for instantaneous seed segmentation, and this Region of Interest (RoI) is then fed into the proposed CNN model for germination prediction in the second stage with an accuracy of 84%. To increase the efficiency a novel seed segmentation and classification model uses U-Net and CNN architectures were used.

The suggested method analyses seed germination using U- Net's segmentation of images and CNN's classification. The suggested fusion model has 0.91 pixel accuracy, 0.84 IoU, and 0.90 precision. Seed Encoding and Decoding Network (SeED-Net) achieves 0.94 accuracy with proposed encoding and decoding segmentation methods. Although the model has considerable analytical capabilities, model weights around 1.8 MB of size are suited for implementation on the Jetson Nano embedded GPU. This technology provides significant advantages to farmers by allowing them to assess and evaluate the germination viability of seeds prior to the planting season.

Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
CNN	Convolutional Neural Network
R-CNN	Region-based Convolutional Neural Network
Mask R-CNN	Mask Region-based Convolutional Neural Network
GAN	Generative Adversarial Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
ReLU	Rectified Linear Unit
NMS	Non-Maximum Suppression
SGD	Stochastic Gradient Descent
RMSprop	Root Mean Square Propagation
Adam	Adaptive Moment Estimation
IoU	Intersection over Union
mAP	Mean Average Precision
FPN	Feature Pyramid Network
ROI	Region of Interest
U-Net	U-Net Convolutional Networks
VGG	Visual Geometry Group
ResNet	Residual Network

YOLO	You Only Look Once
SSD	Single Shot MultiBox Detector
NLP	Natural Language Processing
BERT	Bidirectional Encoder Representations from Transformers
NMT	Neural Machine Translation
FCN	Fully Convolutional Network
GRU	Gated Recurrent Unit
MSE	Mean Squared Error
MAE	Mean Absolute Error
DNN	Deep Neural Network
GANs	Generative Adversarial Networks
BatchNorm	Batch Normalization
FLOPs	Floating Point Operations Per Second
TPU	Tensor Processing Unit
GPU	Graphics Processing Unit
CUDA	Compute Unified Device Architecture
API	Application Programming Interface
ML	Machine Learning
AI	Artificial Intelligence
ROI	Region Of Intrest
RPN	Region Proposal Network
FC	Fully Connected Layer
GPU	Grpahics Processing Unit
EGPU	Embedded Grpahics Processing Unit
SVM	Support Vector Machine

MLP	Multilayer Perceptron (a type of neural network)
LAD	Least Absolute Deviations

Chapter 1

Introduction

In the modern era of agriculture, ensuring the highest quality of seeds is paramount for maximizing crop yield and sustainability. The thesis titled "Design and Development of Seed Germination Classification System using AI and IoT" embarks on innovating seed testing by integrating Artificial Intelligence (AI) [1] and the Internet of Things (IoT). To accurately classify seed germination potential, this approach aims to revolutionize traditional methodologies by offering precision, efficiency, and scalability in assessing seed quality. Seed testing is a critical process that evaluates the quality of seed lots, focusing on aspects such as physical purity, moisture content, germination rates, and overall health. The objective is to eradicate the risks associated with planting low-quality seeds, which can adversely affect agricultural output. The importance of seed testing is underscored by its ability to identify quality problems, determine seeds' suitability for planting, and ensure adherence to established quality standards.

The quality of seeds is assessed based on various attributes: purity (physical and genetic), moisture content, germination, vigor, and freedom from diseases insect infestation. Sampling in seed testing laboratories involves a systematic approach to obtaining representative seed samples. This process is crucial for conducting accurate tests, and it employs methods like mechanical dividing, modified halving, hand halving, the random cup method, and the spoon method to achieve homogeneity and the desired sample size for analysis. Incorporat-

ing AI and IoT into seed germination classification systems presents an innovative solution to automate and refine the seed testing process. This technology will advance real-time data collection, analysis, and decision-making, enhancing the accuracy and reliability of seed quality assessments.

By leveraging AI algorithms, it is possible to predict germination outcomes and classify seeds based on their viability and quality efficiently. Meanwhile, IoT devices facilitate continuous monitoring and data gathering throughout the testing phases, ensuring comprehensive analysis and insights. This thesis endeavors to explore the intersection of technology and agriculture, specifically focusing on improving seed quality evaluation through the development of an AI and IoT-based germination classification system. The integration of these technologies promises to bring about significant improvements in agricultural practices, seed quality control, and ultimately, crop production. Seed is of great importance in agriculture and plays a fundamental role in the development of crops. It is essential for domesticating plants and introducing new cultivars and hybrids into the market [2]. Producing high-quality biological products for public health relies on using suitable seeds based on standard principles [3]. Seed development, including embryo genesis and seed maturation, determines the quality of sown seed, especially under abiotic stress [4]. Seeds are vital for sustainable food and feed production, public health, and the development of agriculture.

The quality of seeds is crucial in agriculture since it has a direct impact on crop yield and the security of food supply. It is influenced by various factors including genetic factors, environmental cues, and growing conditions [5]. The evaluation of seed quality involves assessing parameters such as physical and genetic purity, physiological quality, seed vigour, and health [6]. Seed quality testing should be done following standard procedures to ensure consistent and reproducible results [7]. Enhancing seed quality through various technologies is crucial for optimal crop performance under different environmental conditions and stressors [8]. Overall, seed quality assessment and enhancement play a vital role in achieving sustainable agriculture and meeting the nutritional needs of the growing population.

Seed quality detection methods have been a focus of research to optimize crop establishment and improve agricultural practices. Seed quality evaluation is an essential component

of seed testing laboratories. The accuracy and precision of sampling and testing procedures are crucial for reproducible and reliable results. Measurement uncertainty is evaluated for seed quality indicators, such as thousand-seed weight, in accredited testing laboratories [9]. Physical and physiological quality tests are necessary to estimate the efficacy of seeds, and the deterioration of seeds can affect their quality [10]. Computer vision-based systems can automate the quality assessment process, providing accurate results for physical purity testing of seed samples [11].

1.1 Motivation & Objectives

1.1.1 Motivation

Germination testing of seeds is an important process to determine their viability and quality. Various methods and techniques have been explored to assess germination rates and seed characteristics. Authors conducted a study on turi seeds and found that pre-treatments such as soaking seeds in sulfuric acid improved seed viability [12]. Djurdjic et al. investigated the effects of low temperatures and other abiotic factors on seed germination in winter wheat varieties [13]. Image processing techniques are used in several studies to enhance seed germination testing [14]. Image processing has potential applications in seed germination testing. Digital image processing techniques can accurately predict seed defects by analyzing seed pictures. Furthermore, radiographic image processing can be used to associate seed morphology and tissue integrity with the physiological potential of seeds, allowing for the identification of seeds with low physiological potential [15].

Deep learning models based on convolutional neural networks (CNN) can analyze RGB image data to classify seeds based on germinability and usability, allowing seed companies to avoid wasteful disqualification of seed lots [16]. Additionally, deep learning models can be trained on X-ray images to monitor the quality of seeds based on their internal tissues, germination, and vigor, providing non-destructive and robust classification [17]. These models can accurately discriminate seeds based on their internal tissue integrity, germination,

and vigor, providing relevant information on the physical and physiological parameters of the seeds [18]. Secondly, the amount of data received from sensors in smart farms can be utilized to make informed decisions for optimal planting, land use, yield improvement, and disease/pest control using deep learning models [19]. Additionally, deep learning combined with seed imaging and machine- learning based phenotypic analysis can automate large-scale germination scoring, providing reliable analysis of germination-related traits [20].

Using IoT and deep learning to improve seed germination offers several potential benefits. It allows for automatic monitoring of the germination process, eliminating the need for manual intervention thereby reducing labor and time requirements [21]. It also enables accurate detection and classification of germination status, even in complex situations such as water stains or impurities [22]. Facilitates seed classification based on germinability and usability, allowing seed companies to avoid wasteful disqualification of seed lots and financial losses.

Additionally, the integration of IoT sensor data, high-resolution imagery, and manual intervention data in a synchronized time-series database environment provides a comprehensive view of plant conditions and better means for intervention [23].

Illustrative figure 1.1 presents a multifaceted approach to agricultural innovation, focusing on seed germination studies and the application of AI to enhance crop health and sustainability. At the heart of this initiative lies the existing and proposed growth chamber systems, designed to simulate and monitor the environmental conditions affecting seed germination dynamics. These chambers play a critical role in the in-depth analysis of seed germination processes and the factors influencing them.

1.1.2 Objectives

The research objectives are formed based on the identified gaps in the literature. The summary of the research objectives are given as follows:

Objective 1: To design and implement growth chamber for seed germination dataset using IoT interface.

Objective 2: To predict seed germination using two stage CNN network.

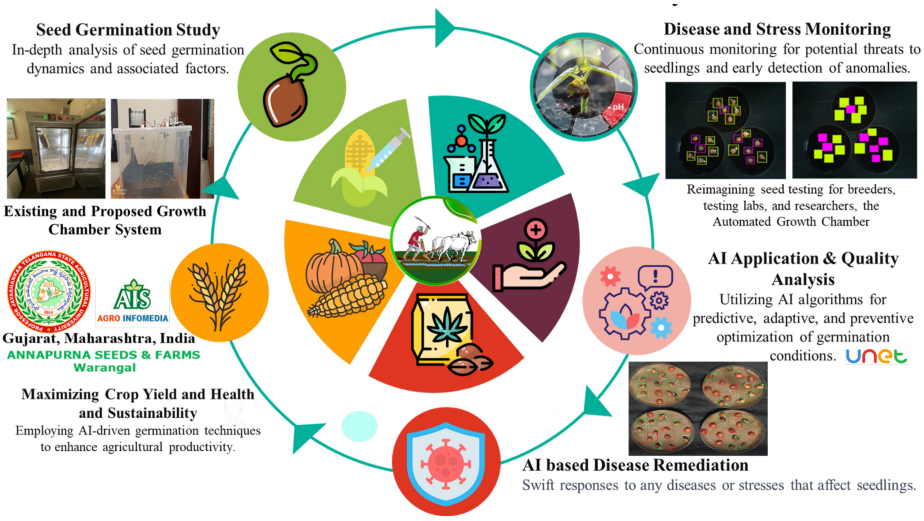


Figure 1.1: Graphical representation showing automated seed germination detection system

Objective 3: To detect seed germination and evaluate using semantic segmentation techniques and enhancing performance metrics.

Objective 4: Deploy a proposed deep learning model on embedded GPU in real time analysis.

1.2 Summary of the Contributions

In this section, an overview of chapter-wise contributions to this thesis is presented. Each subsection presents a summary of the contributions of the corresponding chapter. The thesis consists of seven chapters as follows. The content of each of these chapters is described briefly below:

Chapter 2: The chapter starts by presenting a comprehensive summary of the literature study, which establishes the framework for the future topics. The literature review presents key topics and ideas, providing an overview of the existing knowledge of the study issue. The work discusses noteworthy studies and research results, highlighting the progress achieved in the subject. In addition, the chapter assesses the current level of understanding, highlighting any deficiencies and constraints in the existing body of research that justify the need of the

current study.

Chapter 3: This part of the chapter presents a model that utilizes the capabilities of Mask RCNN, a powerful framework, to accurately identify and classify different types of seeds in images. The use of sophisticated technology to identify different seed varieties is crucial for streamlining the automation of seed quality evaluation and optimizing the environment for germination.

Chapter 4: This chapter presents a two-stage network that uses CNNs to automate seed detection and germination evaluation. In the first step, Mask R-CNN framework segments seeds instantly, and in the second stage, the CNN model predicts germination using this Region of Interest (RoI). This method improves Mask-RCNN performance by applying dual-stage deep learning.

Chapter 5: This contribution proposed a seed detection and germination fusion model. The proposed seed segmentation and classification model uses U-Net and CNN architectures. The proposed method analyses seed germination using U-Net's image segmentation and CNN's classification. By using the Nvidia Jetson Nano embedded GPU platform, the model is optimized for real-time processing and applications. This method improves detectron performance by using semantic segmentation. A jetson nano ported fusion model predicted seed germination. Fusion model improves seed germination accuracy.

Chapter 6: This contribution proposed a deep learning architecture Seed Encoding-Decoding Instant Segmentation Network (SeED-Net) accurately segments and classifies seeds and germination. Pixel-level analysis on seed datasets is done utilizing an encoder-decoder technique and an integrated GPU.

Chapter 7: This section concisely outlines the primary findings and conclusions gained from the study. Furthermore, it offers a concise assessment of potential areas that require further examination. The seed germination experiment using IoT and deep learning has a broad potential. It involves adapting the technology to different crop kinds to improve its integration into smart farming systems for better agricultural decisions. More powerful GPUs will enable real-time analytics by processing data quickly and precisely. Cloud computing will allows huge data to be used for more accurate agricultural predictions, while AI's adaptive learning

will improve models over time.

Chapter 2

Literature Survey

A comprehensive literature review of different works is presented in this chapter. Germination testing of seeds is an important process to determine their viability and quality. Various methods and techniques were explored to assess germination rates and seed characteristics.

The usage of machine vision technology has been extensively implemented across diverse domains of agricultural production [24]. Many studies presented digital image processing for non-destructive rice seed modification, with most rice seed pictures processed utilizing three key aspects: texture [25], color [26], and morphological features [27] [28]. The author has presented an approach that utilizes the Hue Saturation Value (HSV) colour model for the purpose of classifying the germination status of seeds based on the disparities in colour between the seeds and roots. This method was effectively employed in the germination process of seeds with an accuracy of 96.37% [29].

2.1 Handcrafted Features

An automated monitoring system was developed to oversee the complete sunflower seed growth process. The system employed colour, edge, and other relevant data, and was implemented under controlled lighting, temperature, and humidity conditions. This system can serve as a foundation for designing environmental parameters during the breeding process [30]. The methods are predicated upon established image analysis techniques that predomi-

nantly evaluate the germination status through the chromatic attributes of the seed coat and sprout. As a result, their applicability is limited to seed types and requires precise experimental conditions. In addition, the utilization of these methodologies necessitates advanced image acquisition apparatus, thereby restricting their extensive application. In addition, the market offers a limited number of automated solutions specifically developed for the purpose of performing seed germination assays. The germination scan alyzer is a sophisticated automated apparatus that was created in Germany to detect seed germination. The present system employs distinctive blue or grey filter paper as a substrate for the placement of seeds. The process involves the determination of the center of mass through the measurement of seed weight, while the length of germination is ascertained by analyzing the color characteristics of seeds, buds, and the background. Subsequently, this data is utilized to ascertain the state of germination. The efficacy of the Germination Scanalyzer's methodology has been demonstrated to fulfil pragmatic criteria in regard to its consistency and precision [31] . It is important to acknowledge that this approach possesses constraints and distinct utilization prerequisites. The procedure necessitates the use of colour filter paper and entails precise criteria for seed and bud hues, thereby rendering it appropriate for examining solely particular seed varieties.

The Wanshen seed automatic counting instrument, which was domestically developed, is another automated tool used for seed germination testing [32]. This device aids in the automation of the process of counting seeds. In contrast to the previously mentioned approach, this method does not necessitate a prerequisite for the colour of the seed. Rather than manual counting, this method utilises shape-fitting algorithms to achieve precise seed counting automatically. Nonetheless, in the case of this apparatus, it is imperative that the seeds be positioned on a designated plate that emits white light. Although the device is capable of automatically providing the total number of seeds, it lacks the ability to automatically assess the germination status of said seeds. Seed germination testing using image processing faces several challenges. One challenge is the shortage of expertise, time consuming training, and the need for large numbers of reference specimens, which hinders seed identification . Another challenge is the difficulty in segmenting uniform and homogeneous objects from the

background in seed images, which is crucial for extracting morphological and colorimetric features [33].

In contrast to AI and IoT-based systems, which are gaining traction in other agricultural applications for their automation and precision, the existing frameworks for seed germination testing lack significant integration of such advanced technologies. These traditional systems do not capitalize on the potential of machine learning algorithms to learn and improve over time, which could otherwise address the issue of accuracy in non-destructive testing. Nor do they employ IoT devices for continuous, real-time monitoring and data collection at a large scale. Overall, the current landscape of seed germination testing is marked by a gap in the adoption of newer technologies that could lead to more efficient, accurate, and less subjective methods. Moving forward, integrating AI to intelligently analyze images and IoT to facilitate constant data gathering and analysis could revolutionize seed quality testing, making it faster, scalable, and more reliable.

2.2 Deep Learning Features

The agriculture domain [34] lacks explicit reference to the utilization of AI-based autonomous monitoring and predictive tasks [35]. Furthermore, in terms of autonomous operation, the restricted energy storage and elevated power consumption have been a significant area of focus. As previously observed, the integration of computer vision [36], [37], and machine learning [38] has the potential to effectively tackle the issue of seed germination control in the context of industrial automation. Most of the methods do not incorporate an automated data collection model for acquiring various germination states.

In Agriculture, horticulture, and ecological studies, seed germination, the process by which a dormant seed sprouts into a new plant, is a crucial step in plant development [47]. In current agricultural research, there has been an increasing focus on enhancing the evaluation techniques for seed quality through the utilization of sophisticated imaging technologies and machine learning algorithms. The study carried out in this field is noteworthy as it showcases the potential of these methods in improving the precision and effectiveness of seed quality

Table 2.1: Summary of research work done in the area of seed germination prediction.

Year, [ref]	**Methods Used**	**Outcomes**
(2022) [39]	Linear regression, Random Forest, Artificial Neural Networks, REPTree, and M5P methods. Temperature and storage duration as input variables.	Machine learning techniques correctly predicted the quality of stored soybean seeds. Combining temperature and storage time produced the best reliable prediction.
(2021) [40]	Deep learning-based object detection with Faster R-CNN. Development of SeedQuant.	Developed an automatic seed census tool with 94% accuracy. Reduced the time required for counting seeds from 5 minutes to 5 seconds per image.
(2022) [41]	DCNN, LSTM, Attention mechanism.	The attention mechanism-based DCNN-LSTM model performed better. The model's prediction accuracy outperformed alternative approaches by a large margin.
(2023) [42]	Deep learning for rapid estimation of maize seedling emergence using RGB images from UAVs.	With an accuracy of 92% and an average R2 value of 0.96, the system performs well in maize seedling count prediction.
(2022) [43]	X-ray imagery and CNN-based models for predicting tomato seed viability.	Accurate predictions of tomato seed viability are made via CNNs and X-ray imaging. The CNN-based model's accuracy is higher than the image-processing-based model (86.01%).
(2021) [44]	GAN for forecasting future image frames. Developed with PyTorch.	The suggested approach performs well in predicting the segmentation of plant leaves and roots.
(2021) [45]	Deep learning, RGB image data, Word2vec, Domain2vec, GO2vec encodings with RCNN and MLP models.	The research offers a general germination prediction system. Seven vegetable crops' disqualified seed batches were successfully classified.
(2023) [46]	Disinfectants evaluation with MLP, RBF, GRNN networks.	The GRNN algorithm had superior predictive accuracy. The NSGA-II was effective in optimizing disinfectant levels and immersion time.

assessments, thereby resulting in increased crop yields and overall agricultural productivity [48].

Accurate and effective seed germination detection is crucial for a number of applications, including seed quality evaluation [49], agricultural production optimization, and plant growth dynamics monitoring. Traditional techniques for detecting seed germination some- times depend on labor-intensive, inaccurate human counting, or observation. By using convolutional neural networks (CNNs), deep learning (DL) techniques have shown promising results in automating seed germination detection in recent years [50]. Providing crucial knowledge on seed quality and management, covering all areas of seed germination, dormancy, and technology in the field of seed research [51]. Temperature and water potential conditions necessary for the successful germination of various chickpea cultivars are examined, thereby enhancing cultivation practices [52].

To enhance chickpea farming, the appropriate temperature and water potential for germination must be determined [53]. Various applications of DL in agriculture, such as crop monitoring, yield estimate, and precision farming, are examined [54]. The purpose of using CNN is to improve agricultural processes by improving reliability and efficiency. Machine learning uses contemporary artificial neural networks and region suggestions to identify seed germination accurately in high-throughput tests. A machine vision system for seed germination analysis is presented in this work. The hardware configuration and image processing methods utilized to extract pertinent information from seed pictures are described. To distinguish between seeds that have germinated and those that have not, the authors use conventional techniques, such as color analysis, texture analysis, and shape-based characteristics [55]. To differentiate between seeds that have germinated and those that have not, the authors use a CNN-based architecture and form descriptors.

The findings demonstrate how DL techniques are successful at precisely recognizing seed germination phases [56]. A two-stream DL architecture for segmenting and counting leaves from 2-D plant photos encompasses size and shape differences [57]. Different plant phenotype models are discussed in [58] and [59]. Simpler empirical modthat incorporate microclimate and soil parameters may be able to forecast seedling emergence with adequate accuracy

[60]. Real-time strawberry detection using deep neural networks on embedded system, It improves deep neural network detection by revamping the neural structure and decreasing parameters, making it ideal for edge computing [61]. George-Jones et al. [62] automated vestibular schwannoma development detection with U-Net, demonstrating its medical image processing capabilities. Karimi et al. [63] built a pixelwise U-Net model to evaluate wheat planting success in agriculture. Qian et al. [64] also improved the U-Net framework for quantitative melon fruit phenotypic trait analysis, demonstrating its versatility for agricultural applications.

SDDNet is a neural network that has been developed specifically for the purpose of real-time crack segmentation, which is a critical task in the context of industrial infrastructure maintenance [65]. An attention-based generative adversarial network uses thermography to emphasize and identify interior structural defects [66]. The use of a dual encoder–decoder deep model has been shown to improve the accuracy of polyp segmen- tation in colonoscopy images. This helps move colorectal cancer prevention [67]. Res-UNet for classification of tree species using U-Net and ResNet. Res-UNet uses ResNet's feature extraction while keeping U-Net's geographical information by changing the final layers to multiclass classification instead of semantic segmentation [68]. However, for researchers interested in this field of inquiry, there is a lack of publicly accessible benchmark datasets particularly created for investigating the germination of seeds. In contrast to conventional approaches that depend on manual examination and subjective evaluation, the utilization of advanced imaging and machine learning methodologies for seed quality evaluation presents enhanced objectivity and standardization. Consequently, this approach yields heightened precision and consistency in the identification of crucial seed attributes. In addition, the automated characteristics of these contemporary methodologies allow for expedited and more effective examination of seed samples, thereby facilitating extensive assessments to cater to the requirements of contemporary agricultural practices [69].

Chapter 3

AI-Enhanced Seed Quality Assessment with Environmental Control using Mask RCNN

In this chapter the proposed model uses the features of MASK RCNN, a robust framework, to properly recognize and categorize distinct seed types in an image. This sophisticated seed-type identification technology is critical for automating seed quality evaluation and optimizing growing conditions. Our methodology uses a seed image and the MASK-RCNN architecture to detect and classify different kinds of seeds accurately. The suggested approach automatically adjusts the temperature, moisture, humidity, and light based on the type of seed. The model performs an accuracy of 86% compared to state-of-the-art pre-trained models such as VGG-16, Google Inception, and ResNet-50, which are used as the backbone of the proposed MASK RCNN architecture.

Chapter Organization: Section 3.1 presents the proposed methodology. Section 3.2 discusses the experimental settings. Section 3.3 discusses the experimental results and analysis.

3.1 Methodology

In our proposed work,

1. To automate and optimize temperature, humidity, light, and moisture for optimal germination.

2. To develop an AI-growth chamber using deep learning to recognize varied seed varieties.

The proposed Mask RCNN architecture model is shown in figure 3.1. The 2-level proposed model is described in this section.

Initially, the data set will be pre-processed using IQR. Then, the dataset underwent further experiments to determine the best oversampling technique for class balance.

3.1.1 Architecture of the Proposed Model

The Mask R-CNN [70] architecture is built for object identification and instance segmentation, and its use in seed detection illustrates its accuracy is shown in figure 3.1. Mask R-CNN is built on a pre-trained convolutional neural network which is essential for feature extraction. This backbone network uses a sequence of convolutions and pooling operations to capture different degrees of information, ranging from simple edges and textures in the early layers to more sophisticated patterns in the deeper layers. This network is good at catching subtleties in seed forms, sizes, and textures in the context of seed identification, laying the groundwork for subsequent research. The RPN (Region Proposal Network) [71] lies above the backbone network. This network is critical in analysing the backbone's feature map and suggesting possible object areas or bounding boxes known as ROIs (Region Of Interest).

The RPN is critical in finding probable seed sites within an input image for seed identification. It creates various area suggestions with varying sizes and aspect ratios and classifies them as seed or no-seed, while also revising their coordinates for greater precision. The addition of the ROI Align layer to Mask R-CNN over its predecessors is a significant improvement over its predecessors. ROI Align, as opposed to ROI Pooling, employs bilinear

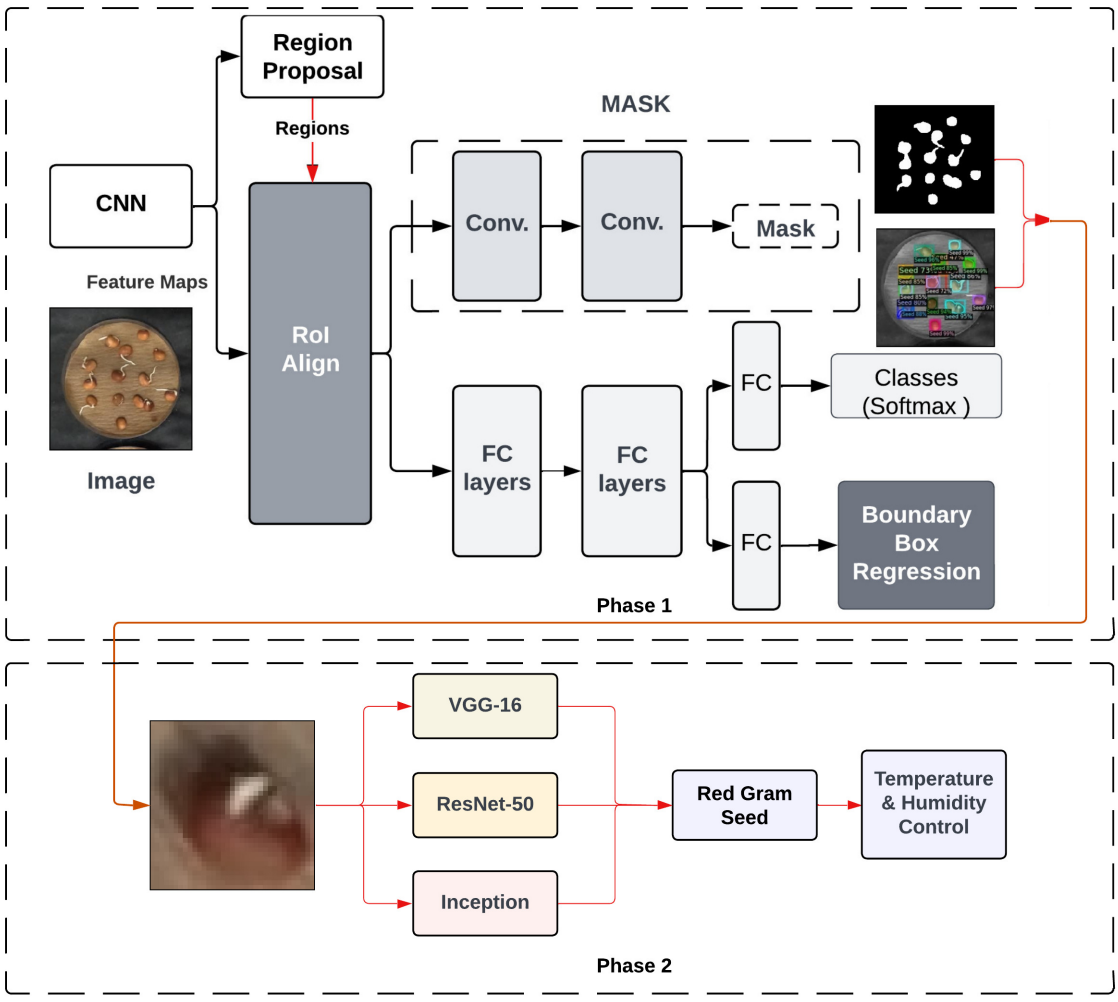


Figure 3.1: Architecture of the proposed Mask-RCNN for the type of seed and environmental parameters monitoring

interpolation to properly map the extracted features to the ROI, correcting misalignments and guaranteeing that the features match exactly to the original picture. This accuracy is critical in seed detection to guarantee that each seed's characteristics are precisely recorded, which is required for accurate classification and mask prediction. In addition, fully linked layers for classification and bounding box regression are included in the network. These layers take the perfectly aligned ROI features and execute two crucial tasks with them. They begin by classifying each ROI and identifying whether it contains a seed. Second, they adjust the bounding box coordinates to snugly encapsulate the identified seed. This phase is essential for correctly

identifying and delineating each seed inside a picture. Mask R-CNN's notable feature, particularly relevant to seed identification, is its mask prediction branch. This branch generates a high-resolution binary mask for each class in each ROI using a convolutional neural network that operates on the ROI-aligned features. However, only the mask associated with the projected class is considered. This corresponds to thorough, pixel-level segmentation of each seed in seed detection, a vital skill for accurate seed analysis and classification.

The architecture of the proposed model is shown in table 3.1. The proposed model's initial phase is feature extraction, a Conv2D layer (3x3, stride 1, output 224x224x64). This layer filters the input picture with 64 distinct 3x3 filters, successfully capturing fundamental elements such as edges and textures. This primary layer is critical for seed recognition because it begins the process of separating seeds from the background and determining fundamental seed features. The stride of 1 maintains the original spatial resolution, guaranteeing that no important information is lost at this step. A convolutional layer calculates output Z for each filter by convolutioning input image X with filter W and a bias term b. One output pixel is shown in equation 3.1.

$$Z_{ijk} = \sum_{m,n} X_{(i+m)(j+n)k}.W_{mnk} + b_k \tag{3.1}$$

3.1.1.1 Regional Proposal Network (RPN)

The RPN is a critical component of Mask R-CNN, responsible for creating high-quality item proposals. RPN functions as a compact FCN (Fully Convolutional Network) that accurately forecasts possible object areas in an image. The input is a shared convolutional feature map, usually obtained from a pre-trained model such as ResNet, and the system carries out two main tasks that is prediction of objectness score and the other is bounding box regression.

3.1.1.2 Predicting the Objectness Score

RPN uses a binary classifier to forecast the probability that each prospective zone will have an object. The equation expresses this mathematically in equation 3.2.

$$p_{obj}(x) = (f_{cls}(x)) \tag{3.2}$$

The equation 3.3 calculates the probability of an object being present at position x by applying the sigmoid function to the output of the classification function at x.

$$t_i(x) = t_i^{reg}(f_{reg}(x)) \tag{3.3}$$

where the function $p_{obj}(x)$ denotes the probability of an item being present in an area located at x. The symbol represents the sigmoid activation function, which transforms input values into a range between 0 and 1. The function $f_{cls}(x)$ represents the output of the classifier branch of the RPN at point x.

3.1.1.3 Anchor Boxes

RPN uses a predetermined set of anchor boxes with variable scales and aspect ratios to detect seeds of varying sizes and forms. Anchor boxes are utilised throughout the whole feature map to enable the network to learn objectness scores and bounding box offsets for each position. The function in the Mask R-CNN architecture is a Shared Feature Extraction. RPN utilises convolutional features that are also used in later stages of Mask R-CNN, which helps enhance performance by preventing repetitive feature calculations. RPN creates object proposals and assigns objectness scores to them. Further stages of Mask R-CNN utilize the ideas as input for performing classification and mask segmentation tasks. Because of shared feature extraction, RPN is notably faster than, enhancing the efficiency of Mask R-CNN. RPN improves the selection of top-tier object regions, resulting in enhanced performance for object detection and mask segmentation tasks. Ultimately, the RPN is crucial in Mask R-CNN for effectively producing object proposals using shared feature extraction and anchor boxes. This enhances the model's overall accuracy and efficiency in seed detection and precise instance segmentation tasks.

3.1.1.4 Bounding Box Regression

RPN simultaneously adjusts the placement of predefined anchor boxes by regressing their coordinates to align them more accurately with the real object boundaries. Mathematically representation of the regression is shown in equation 3.4.

$$t_i(x) = t_i^{reg}(f_{reg}(x)) \tag{3.4}$$

$t_{i(x)}$ is the forecasted bounding box adjustment for the i-th anchor box at position x.

$t_i^{reg}(x)$ is the result of the regression branch of the RPN for the anchor box at the i^{th} position x.

3.1.1.5 Mask Branch

Anchor boxes are utilised throughout the whole feature map to enable the network to learn objectness scores and bounding box offsets for each position of seed.

The Mask branch of Mask R-CNN refines seed proposals generated by the RPN into pixel-wise masks, properly delineating the contours of particular seeds in an image. The RPN sends ROIs to the Mask Branch. The ROIs represent possible seed positions together with their respective bounding boxes. Convolutional processes can cause the ROIs to not line up properly with the feature map grid. This misalignment may result in imprecise feature extraction. To tackle this difficulty, the Mask Branch uses ROI Align, which is a critical step after the RPN's output. ROI Align makes sure that ROIs and the feature map are in the same place by splitting each ROI into smaller areas and using bilinear interpolation to get features from within each area.

3.1.2 Convolutional Layers

The convolution layer uses a filter to extract the features of the input seed image. Performing multiple filter applications on the image generates a feature map of seed images that aids in the classification of the seed image. For expediency, a 2D input image with normalised pixels

is used.

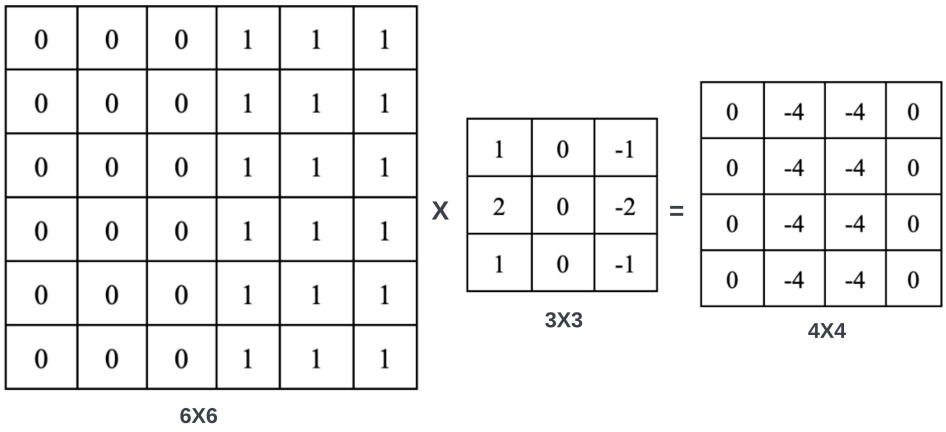


Figure 3.2: Convolution of 6x6 seed pixel image with 3x3 feature detection filter

The 3x3 feature detection filter processed a 6x6 input image. While a single filter was used in this case, it is common practice to employ multiple filters to extract information from seed images. Applying the filter to the input image generates a 4x4 feature map containing information about the image. It generates numerous feature maps. In the first stage, the filter was applied to the highlighted image. Multiplying the pixel values of the image by the filter values and subsequently summing them yields the final value, as depicted in the figure 3.2. As shown in the figure 3.3, move the filter by one column. In this instance, we denote the transition from one column or row to the next as a "stride." At this location, we use a stride of 1 to indicate a column shift of one.

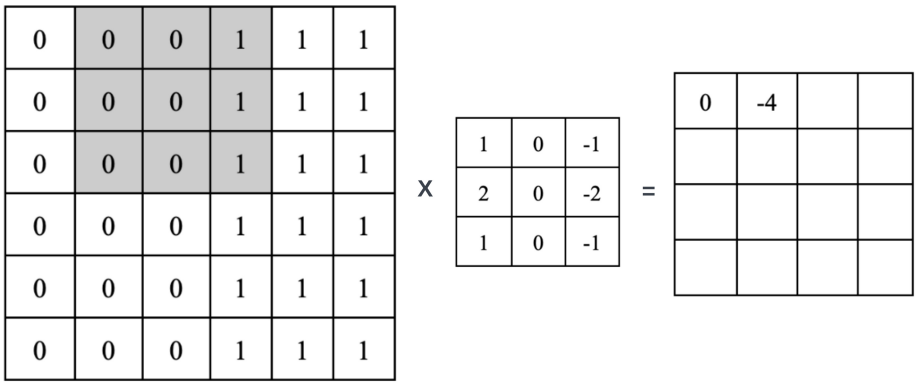


Figure 3.3: Illustrating the movement of 3x3 filter by one

The final feature map is the consequence of the filter traversing the entire image. An activation function introduces non-linearity into the feature map. It is critical to acknowledge that the feature map acquired is proportional to the dimensions of the image. A decrease in the feature map's dimensions is the consequence of augmenting the stride value.

3.1.2.1 Pooling

After the convolutional layer, the pooling layer cuts down the size of the feature map. This helps to keep important information in the input seed image while also accelerating the computation. Pooling produces a variant of the input with a reduced resolution that preserves the essential characteristics of the original seed image. The two most common types of pooling are maximum pooling and average pooling. The figure 3.4 shows the methodology employed in max pooling. Pooling is executed using the feature map acquired from the previous illustration. We are making use of a pooling layer that has a stride of 2 and dimensions of 2x2. Selecting the maximum value from each highlighted region generates a new rendition of the input image with dimensions of 2x2. Pooling reduces the feature map's dimensions.

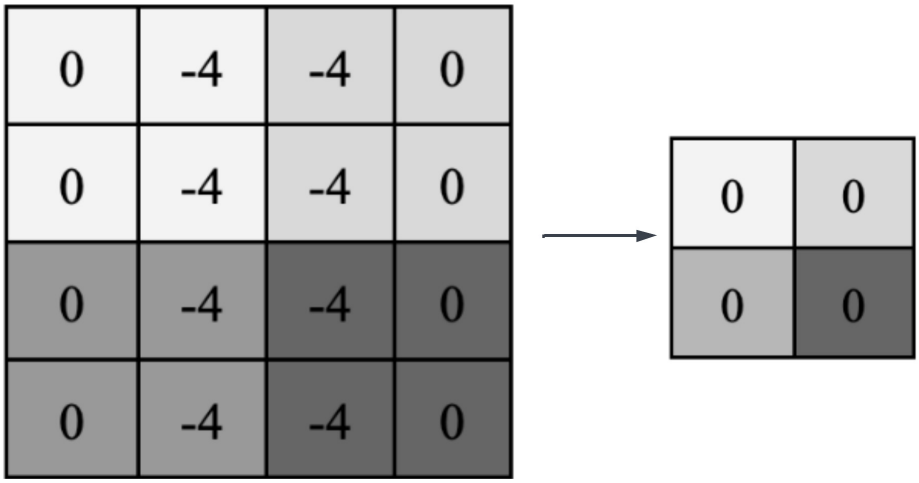


Figure 3.4: Performing maxpooling on seed image

3.1.2.2 Batch Normalisation

BatchNorm is carried out after the Conv2D layer as shown in equation 3.5 and 3.6. This layer standardizes the inputs to the next layer, which is critical for preserving numerical stability and speeding up the training process. BatchNorm guarantees that the network's learning process is efficient and effectiveness of network's learning process by normalizing the output of the preceding layer, which is especially crucial for complicated tasks like seed type categorization.

$$X_{ijk} = \frac{X_{ijk} - \mu_k}{{\sigma_k}^2 + \epsilon} \tag{3.5}$$

$$Y_{ijk} = \gamma_k X_{ijk} + \beta_k \tag{3.6}$$

Following that, the ReLU (Rectified Linear Unit) activation function is used in equaltion 3.7, which introduces non-linearity into the model. This phase is critical as it enables the network to recognize complicated patterns in data that linear models cannot. In the case of seed detection, ReLU assists the network in distinguishing between varied seed shapes, sizes, and textures, which is crucial for correct classification.

$$A_{ijk} = max(0, X_{ijk}) \tag{3.7}$$

Following that, the design includes a MaxPooling2D layer (2x2, stride 2, output 112x112x64). This layer reduces the feature map's spatial dimensions in half, essentially summarising the characteristics in each area. MaxPooling aids in seed recognition by concentrating on more relevant characteristics and lowering the influence of small fluctuations and noise in the picture.The process of MaxPooling involves the application of a 2x2 window, where the resulting output for each window is determined by selecting the maximum value contained inside that window. In the context of a feature map, consider a window that starts at point (i, j) shown in equation 3.8.

$$P_{ij} \in max_{m,n[0,1]} Z_{(i+m)(j+n)} \tag{3.8}$$

Additional Conv2D, BatchNorm, and ReLU layers are added to the network, gradually increasing the depth of the feature map. These layers enable the model to capture more complicated characteristics. As the network's depth rises, it begins to recognize more particular properties of seeds, such as distinct patterns or textures that distinguish one sort of seed from another. The feature maps get increasingly abstract in the latter stages, following multiple cycles of Conv2D, BatchNorm, and ReLU, followed by MaxPooling. The network now has a better knowledge of the many characteristics that distinguish distinct seed kinds. These higher-level characteristics are critical for differentiating between different seed types, particularly when the variations are minor. The network is prepared for classification by the last set of Conv2D, BatchNorm, and ReLU layers, followed by the final MaxPooling2D layer (2x2, stride 2, output 7x7x512). The spatial dimensions are further lowered, and the feature maps are rich in the information required for effectively categorizing various kinds of seeds.

The network then moves to a flattened layer, which turns the 3D feature map into a 1D vector. This transformation is required to feed the convolutional features into fully connected (Dense) layers that oversee classification. The network's ultimate tiers include the Dense and Dropout layers. Dense layers make judgments based on the characteristics collected by the convolutional layers, categorizing the seeds. The Dropout layers are used for regularisation, randomly removing a fraction of the connections (neurons) during training to avoid overfitting and ensure that the model generalizes effectively to new, previously unseen pictures of seeds.

The process of flattening involves transforming a three-dimensional tensor into a one-dimensional vector. If the tensor possesses dimensions. The length of the flattened vector is equal to the product of the dimensions h, w, and d, denoted as h×w×d is shown in equation 3.9.

$$Y(i) = \sum_{j} W_{ij} X_j + b_i \tag{3.9}$$

3.1.2.3 Common Feature Extraction

The Mask Branch enhances the RPN's efficiency by utilizing the shared convolutional feature map obtained from the backbone network (ResNet) used by the RPN. The shared feature map contains a wealth of semantic information about the entire seed image, offering crucial context for tasks like seed classification and mask segmentation. Branching for prediction, Distinct branches process the common feature map for additional processing, similar to the RPN. Classification Branch, This branch, like the one in the RPN, uses fully connected layers to categorize each ROI, estimating the likelihood that each item seed class will be present in the ROI. The class probability for the c-th class as shown in equation 3.10.

$$p(c|x) = softmax(f_{cls}(x)) \tag{3.10}$$

Calculate the probability of class C given input x by applying the softmax function to the output of the classifier function $f_{cls}(x)$.

In this context, the symbols retain their original meaning as previously described. Mask branch uses many convolutional layers and a deconvolutional layer (transposed convolution) to predict a binary mask for each ROI. The binary mask and the ROI have matching spatial dimensions, where each pixel value indicates the probability of that pixel being part of the foreground item located within the ROI. Mathematically representing the expected mask for the ROI of x is as follows 3.11:

$$M(x) = sigmoid(f_{mask}(x)) \tag{3.11}$$

3.1.2.4 RoI Extraction

It enhances mask segmentation accuracy by matching the ROI's with the feature map to ensure precise feature extraction. This expands on the RPN's function in creating top-notch proposals. Shared feature extraction enhances efficiency by preventing repetitive feature calculations, complementing the RPN's role in computational efficiency. The Classification Branch assigns seeds to the RoI class labels. This is necessary to tell the difference be-

tween object seed and no-seed categories and improve the RPN's ability to predict objectness scores. Mask Branch produces masks at the pixel level that precisely delineate the outlines of individual objects within the ROIs, allowing for precise instance segmentation based on the initial object proposals from the RPN. Non-Maximal Suppression (NMS) can be used to eliminate duplicate or overlapping masks, guaranteeing that each item is depicted by a single, top-notch mask.

Finally, the network includes a Dense layer and a Softmax activation function. The Dense layer generates a vector representing the number of seed types, which the Softmax function translates into a probability distribution. This distribution indicates the seed's probability of belonging to each available category, resulting in a clear and probabilistic categorization. The aforementioned function transforms the logits present in Y into corresponding probabilities, as shown in equation 3.12.

$$S(Y)_i = \frac{e^{Y_i}}{\sum_j e^{Y_j}} \tag{3.12}$$

3.1.2.5 Loss Functuion

During training, we use a multi-task loss function is used that combines a classification loss (like cross-entropy) for the classification branch and a binary segmentation loss (such as binary cross-entropy) for the mask branch. This loss function helps the model learn the best parameters for both tasks, similar to like how the RPN uses its own loss function. Ultimately, the Mask Branch smoothly incorporates the RPN, inheriting the task of object suggestions and enhancing them into accurate object masks. The Mask Branch enhances Mask R-CNN by utilising ROI Align, shared feature extraction, and specialised branches for classification and mask prediction. This allows the model to not only identify objects but also precisely outline their shapes in an image, providing a thorough comprehension of the visual environment. The multi-task loss function in training Mask R-CNN optimises the model for object classification and mask segmentation tasks simultaneously. This function usually merges two separate loss components.

Table 3.1: Summary of the proposed Mask R-CNN architecture

Layer	Type	Kernel Size	Stride	Output Size
1	Conv2D	3x3	1	224x224x64
2	BatchNorm	–	–	224x224x64
3	ReLU	–	–	224x224x64
4	MaxPooling2D	2x2	2	112x112x64
5	Conv2D	3x3	1	112x112x128
6	BatchNorm	–	–	112x112x128
7	ReLU	–	–	112x112x128
8	MaxPooling2D	2x2	2	56x56x128
9	Conv2D	3x3	1	56x56x256
10	BatchNorm	–	–	56x56x256
11	ReLU	–	–	56x56x256
12	Conv2D	3x3	1	56x56x256
13	BatchNorm	–	–	56x56x256
14	ReLU	–	–	56x56x256
15	MaxPooling2D	2x2	2	28x28x256
16	Conv2D	3x3	1	28x28x512
17	BatchNorm	–	–	28x28x512
18	ReLU	–	–	28x28x512
19	Conv2D	3x3	1	28x28x512
20	BatchNorm	–	–	28x28x512
21	ReLU	–	–	28x28x512
22	MaxPooling2D	2x2	2	14x14x512
23	Conv2D	3x3	1	14x14x512
24	BatchNorm	–	–	14x14x512
25	ReLU	–	–	14x14x512
26	Conv2D	3x3	1	14x14x512
27	BatchNorm	–	–	14x14x512
28	ReLU	–	–	14x14x512
29	MaxPooling2D	2x2	2	7x7x512
30	Flatten	–	–	1x25088
31	Dense	–	–	1x4096
32	Dropout	–	–	1x4096
33	Dense	–	–	1x4096
34	Dropout	–	–	1x4096
35	Dense	–	–	–
36	Softmax Activation	–	–	1x1000

The model penalizes inaccurate class predictions within the ROIs generated by the RPN. One frequently used option is the cross-entropy loss, which is defined mathematically in equation 3.13

$$L_{cls} = -_{i=1}^{N_c} y_i log(p_i) \tag{3.13}$$

The equation 3.13 calculates the cross-entropy loss. The term L_cls represents the classification loss. The symbol N_c denotes the total number of object classes. The variable y_i represents the true label (0 or 1) for the i-th class, indicating if the class is present in the ROI or not. p_i is the forecasted probability of the i-th class by the classification branch of the Mask Branch for the respective ROI.

The model is penalised for mistakes in the anticipated binary masks created by the Mask Branch. One popular option is the binary cross-entropy loss, represented mathematically as equation 3.14

$$L_{mask} = -x_{x,y} y_{xy} log(p_{xy}) + (1 - y_{xy}) log(1 - p_{xy}) \tag{3.14}$$

The equation 3.14 calculates the cross-entropy loss for a binary classification task. The variable L_{mask} represents the loss associated with the mask. x and y denote the pixel locations inside the ROI mask. The variable y_{xy}represents the true label (0 or 1) for the pixel located at coordinates (x, y), indicating whether the pixel is part of seed or not. The symbol p_{xy} denotes the forecasted likelihood that the pixel at coordinates (x, y) is part of the foreground item produced by the mask branch for the respective ROI.

$$L = L_{cls} + *L_{mask} \tag{3.15}$$

The equation 3.15 expresses the comprehensive multi-task loss function, integrating classification and mask losses. L equals the sum of L_{cls} and the product of lambda and L_{mask}. The symbol "L" represents the total multi-task loss. The symbol denotes a hyperparameter that controls the relative significance of the two separate losses in the training process.

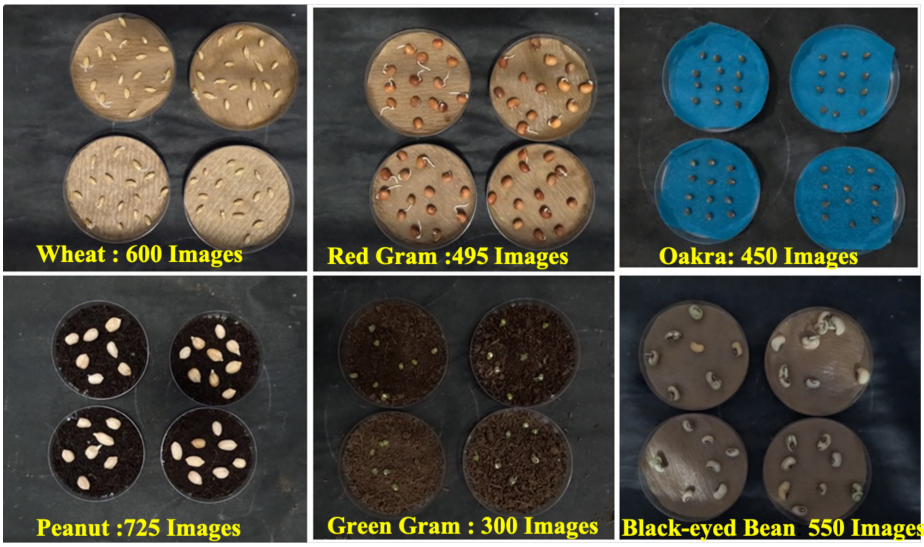


Figure 3.5: Dataset of different varieties of seeds collected using AI growth chamber

3.2 Experimental Results

3.2.1 Dataset Collection

The chamber's constant monitoring allows for the gathering of substantial data on the impact of various environmental variables on seed development. Such information is crucial in identifying seed quality since high-quality seeds often exhibit high germination rates, strong development, and resistance to environmental stresses. Furthermore, the system's capacity to identify early symptoms of abnormalities or aberrant development, allows for appropriate intervention. This AI growth chamber has several benefits. Its combination of precise environmental controls and powerful AI image processing guarantees a very accurate evaluation of seed quality. The system's automation increases efficiency while decreasing the possibility of human mistakes. The extensive data analysis gives greater insights into seed development patterns and the variables impacting seed quality. The chamber's flexibility, allows it to test a broad range of seed varieties as shown in figure 3.5, making it a useful tool for a wide range of agricultural applications. Furthermore, the capacity to recognize issues early in the germination or development phase increases the success rate of seed culture.

The table 3.2 shows the distribution of seed images used for training and testing. Total

2496 seed images have been assigned for training and 624 for testing, for a total dataset size of 3120 seed images.

Table 3.2: Dataset of seed germination images.

Seed Type	Number of Images
Wheat	600
Red Gram	495
Okra	450
Peanut	725
Green Gram	300
Black-eyed Bean	550
Total	3120

3.2.2 Data Annotation

During the annotation process, it is necessary to assign labels to each image frame that depicts a seed with in a petri dish in order to identify the specific seed. The execution of this task is accomplished by employing a software application referred to as RoboFlow [72]. The process is illustrated in figure 3.6.

The first part of the figure, designated as Figure 3.6 (a), displays the "Input Image." This photograph provides a clear, unaltered view of the seed's environment, offering insights into its natural state and surrounding context within the petri dish.

The subsequent section, 3.6 (b), reveals the "Binary mask Image". Unlike its counterpart, this image undergoes digital processing to isolate the seed, presenting it against a stark, uniform backdrop. This stark contrast is instrumental in delineating the seed's contours, significantly simplifying the task of annotation by clarifying which areas require labelling.

3.2.3 Experimental Setup

The AI growth chamber is a novel device in seed quality assessment, integrating modern technology such as a high-resolution camera, accurate temperature and humidity control systems, adaptive grow lighting, a moisture sensor, and a Jetson Nano board running Mask R-CNN. The camera, which records precise images of seeds, is at the core of this system as shown in

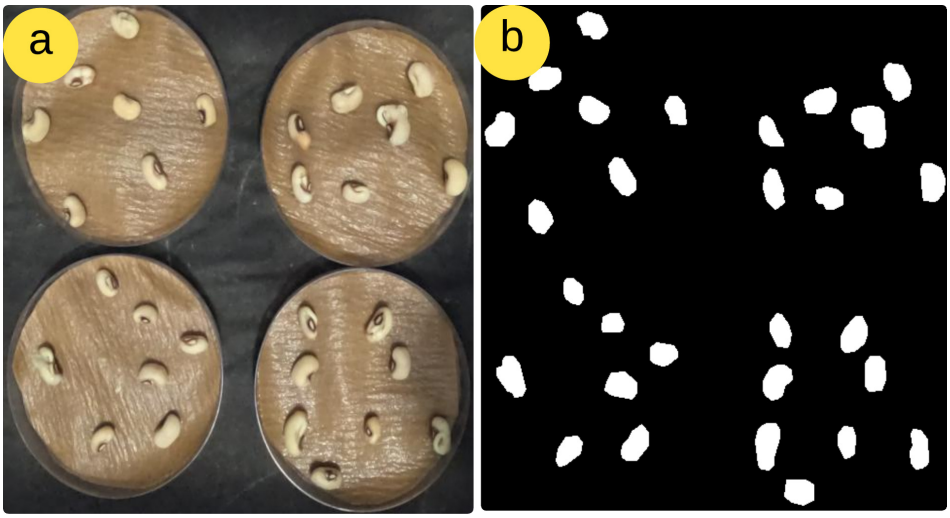


Figure 3.6: Seed dataset (a). Images of the collected dataset (b). Mask images of the seeds

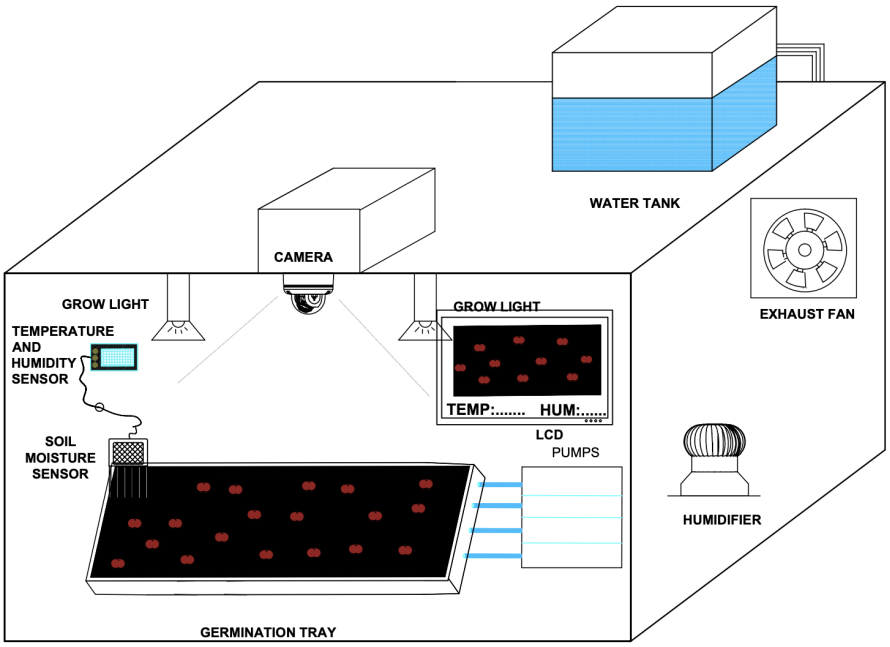


Figure 3.7: Block diagram of the proposed AI-based growth chamber

figure 3.7. These images are processed using the Mask R-CNN algorithm on the Jetson Nano board, a device well-suited for performing complicated computational tasks. The algorithm is crucial in recognizing and categorizing each seed based on its kind. This identification is critical since various seeds have distinct needs for optimum development, which the chamber is meant to offer. The temperature and humidity inside the chamber are rigorously regulated,

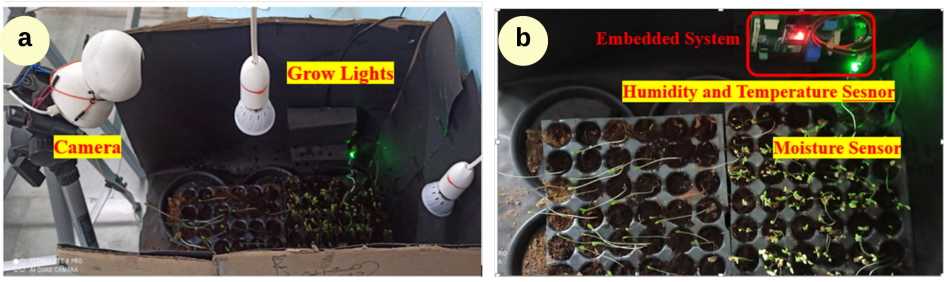


Figure 3.8: Prototype developed for AI-based growth chamber

producing an optimum environment for seed germination and development. This degree of control is required because seeds from various species typically grow in certain climatic circumstances. The chamber as shown in figures 3.8 and 3.9 guarantees reliable evaluations of seed quality by altering these environmental conditions to fit the demands of each seed type. The grow lights inside the chamber imitate natural sunshine, giving the spectrum required for plant photosynthesis. The ability to vary the intensity and duration of these lights is critical for simulating diverse ambient light situations and studying how seedlings react to varied lighting circumstances. The moisture sensor, which constantly checks the water content of the soil, is an important component of the chamber. Proper watering is essential for seed germination, and this sensor guarantees that the seeds get the necessary quantity of water for effective development. After the seeds are placed inside the chamber, the quality testing procedure starts. The Mask R-CNN algorithm finds and classifies them, enabling the chamber's environmental controls to change appropriately. The camera continues to play its function by capturing photographs of the seeds as they germinate,grow, and the algorithm analyzes these images to track their progress. This examination involves determining the germination rate and the health of the developing plants, both of which are important markers of seed quality.

This section describes the datasets, noise levels, and evaluation metrics employed to assess the proposed model. The qualitative and quantitative results of the proposed model are then compared with the existing models.

In order to assess the efficacy of the proposed model in predicting seed types, a set of six widely used measures are employed: Pixel accuracy 3.16, Intersection over Union (IoU) 3.17, recall 3.18, and the F1 score 3.19. The metric of pixel accuracy quantifies the propor-

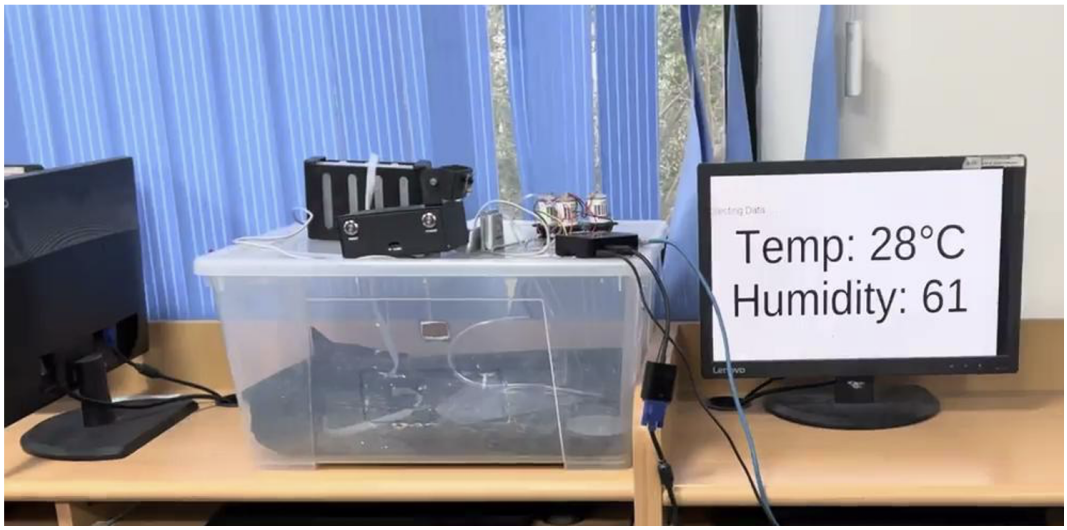


Figure 3.9: AI-Based growth chamber

tion of correctly classified pixels inside an image segmentation task. The Intersection over Union (IoU) metric is used to measure the degree of overlap between the anticipated masks and the ground truth masks. Precision is a metric that quantifies the ratio of properly recognized masks to the total number of masks predicted by the model. On the other hand, recall measures the ratio of correctly identified masks to the total number of masks in the actual ground truth. The F1 score, however, represents the harmonic mean of accuracy and recall, providing a balanced evaluation of these two criteria. The formulas for these measures are presented here to provide a thorough assessment of the network's efficacy.

$$Pixelaccuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{3.16}$$

$$IoU = \frac{TP}{TP + FP} \tag{3.17}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.18}$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \tag{3.19}$$

3.3 Experimental Results

In this section, a comprehensive analysis of the results obtained during the evolutionary process is presented.

The proposed model is trained for 50 epochs on a partitioned dataset before being tested. The dataset is partitioned into 80 and 20 ratios, and the training set is used for Mask RCNN. Figure 3.10 shows the proposed model's output semantically segmented seeds from the petri dish, whereas figure 3.11 shows the individual binary masks created by the Mask RCNN model for the specified petri dish. The binary mask of the input picture is produced by combining these distinct binary masks. As seen in figure 3.12, these segmented pictures are then used to extract seeds. This technique involves creating precise, pixel-level masks for each seed in an image, effectively isolating them from the rest of the visual content, typically against the backdrop of a petri dish. As shown in figure 3.13, after segmenting the seeds, the segmented images serve as a data source for extracting the seeds' characteristics. These extracted features are then used as input for a proposed classifier. These masks highlight the model's ability to accurately identify and delineate various objects within an image. By applying the MASK-RCNN technique, we have successfully demonstrated the model's precision in segmenting objects from their background, showcasing its potential for applications in seed recognition and analysis.

3.3.1 Comparative Analysis

In the domain of seed classification, the ResNet50, Inception, and VGG16 models have unique performance attributes. Evaluation metrics of the classification model with different pretrained models is shown in table 3.3

3.3.1.1 ResNet-50

ResNet50 has notable performance with a pixel accuracy of 0.87, denoting that 86% of pixels were accurately categorized as shown in figure 3.14. This high level of accuracy is crucial in the identification of diverse seed kinds. The Intersection over Union (IoU) value of 0.72

Figure 3.10: Extracting the seeds using proposed model using Mask RCNN

indicates a significant correspondence between the anticipated and actual seed regions, hence demonstrating a dependable segmentation. The model achieved an accuracy score of 0.82, indicating that it correctly recognized seeds 82% of the time. Furthermore, it exhibited a recall score of 0.91, indicating its efficacy in recognizing a majority of seed types. The F1 score, which measures the harmonic mean of precision and recall, demonstrates a strong and reliable performance across all evaluated metrics.

3.3.1.2 Google Inception

It demonstrates a pixel accuracy of 0.83, which suggests a commendable performance in the identification of seed types. A segmentation accuracy of 0.68, as measured by the IoU, indicates a satisfactory level of performance. The precision value of 0.77 indicates that the prediction model accurately classified seeds with an accuracy rate of 77%. Similarly, the recall rate of 0.88 signifies that the model effectively recognized the majority of seed kinds. The F1 score of 0.82 indicates a well-balanced performance, however, there may be rare

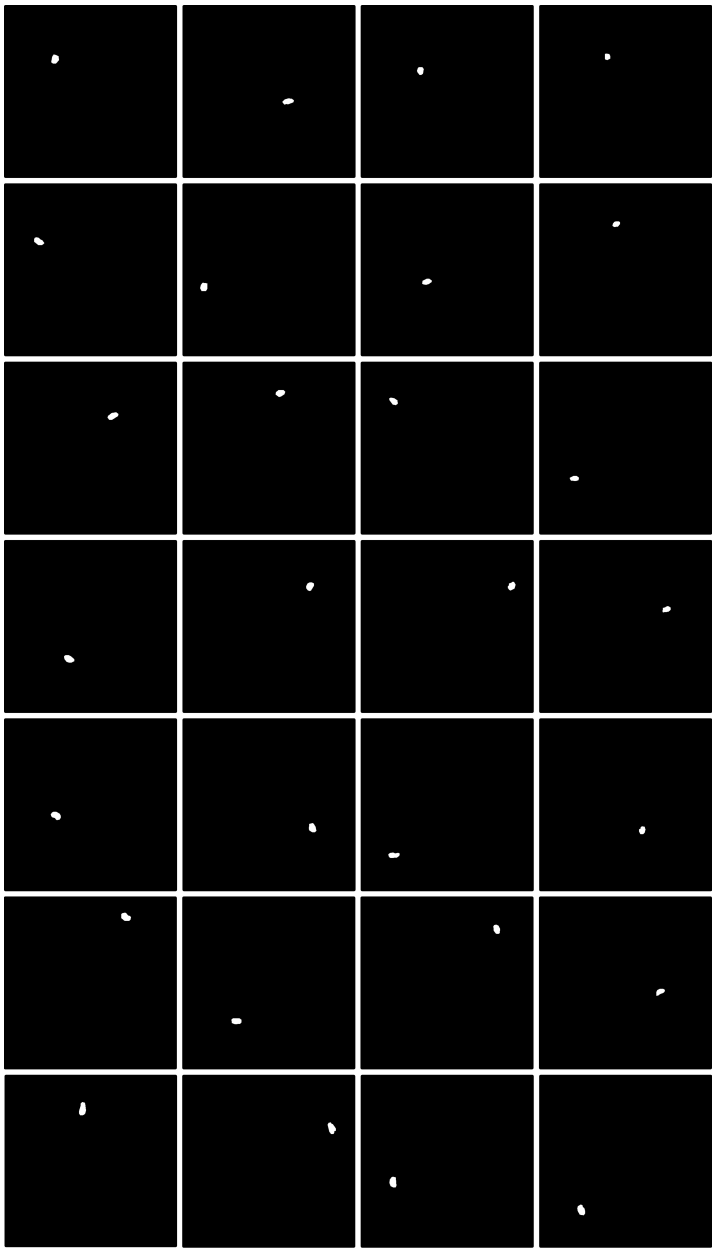


Figure 3.11: Binary masks generated by the proposed model

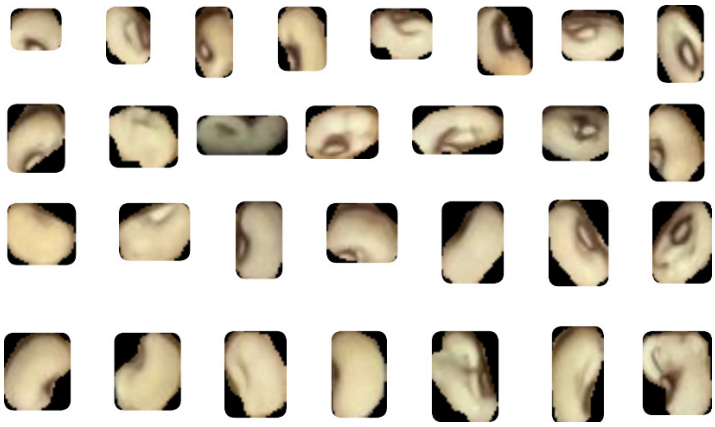


Figure 3.12: Semantic segmented seeds extracted from petri dish using proposed model

Table 3.3: Evaluation metrics of the classification model with different pretrained models

Model	Pixel Accuracy	IoU	Precision	Recall	F1
Mask RCNN + ResNet50	**0.86**	0.72	0.82	**0.91**	0.85
Mask RCNN + Inception	0.83	0.68	0.77	**0.88**	0.82
Mask RCNN + VGG-16	0.84	0.73	0.80	**0.92**	0.86

errors in seed identification. The training and validation loss of Google Inception is shown in 3.15.

3.3.1.3 VGG16

The VGG-16 model demonstrates a pixel accuracy of 0.84, indicating its proficiency in accurately recognizing the majority of seed types and properly categorizing a significant proportion of picture pixels. The IoU score of 0.73 indicates strong segmentation ability. The precision of the model is 0.80, suggesting a high level of accuracy in its predictions with a lower occurrence of false positives. Additionally, the model has a notable recall value of 0.92, demonstrating its effectiveness in identifying different types of seeds. The F1 score, which is 0.86 in this case, demonstrates a strong balance between precision and recall. This characteristic makes it a reliable option for the categorization of seeds. The training loss and validation loss of VGG16 is shown in figure 3.16.

Each model has distinct strengths, with ResNet50 demonstrating a commendable equilibrium across all measures, Inception effectively balancing accuracy and F1 score, and VGG16

Figure 3.13: Masks generated using proposed MASK-RCNN model

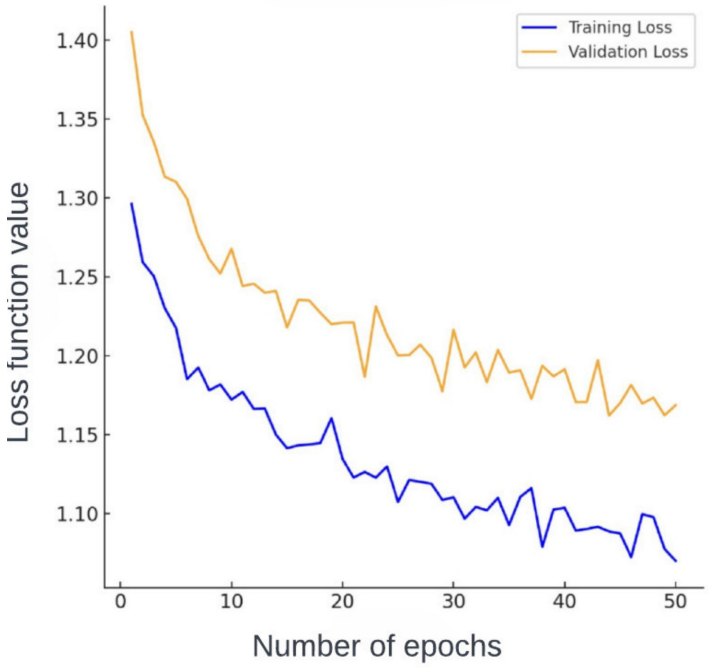


Figure 3.14: ResNet50's model training and validation.

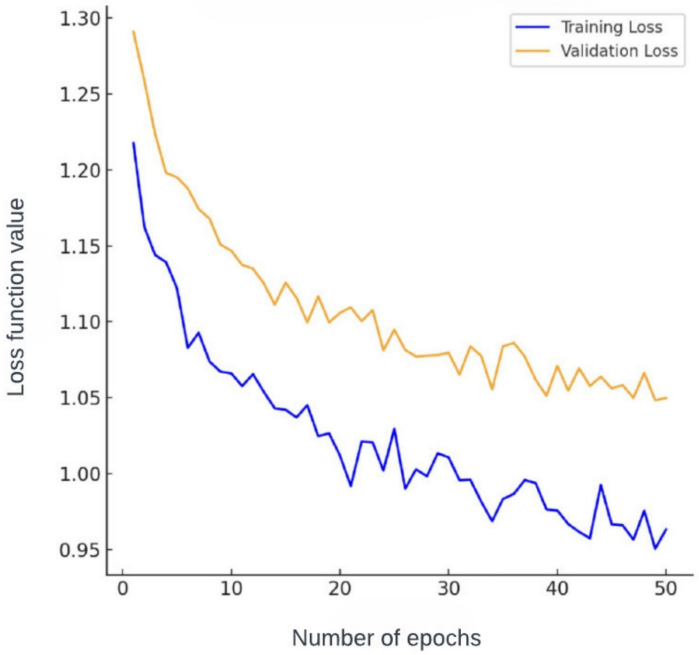


Figure 3.15: Proposed model with Google Inception training and validation losses.

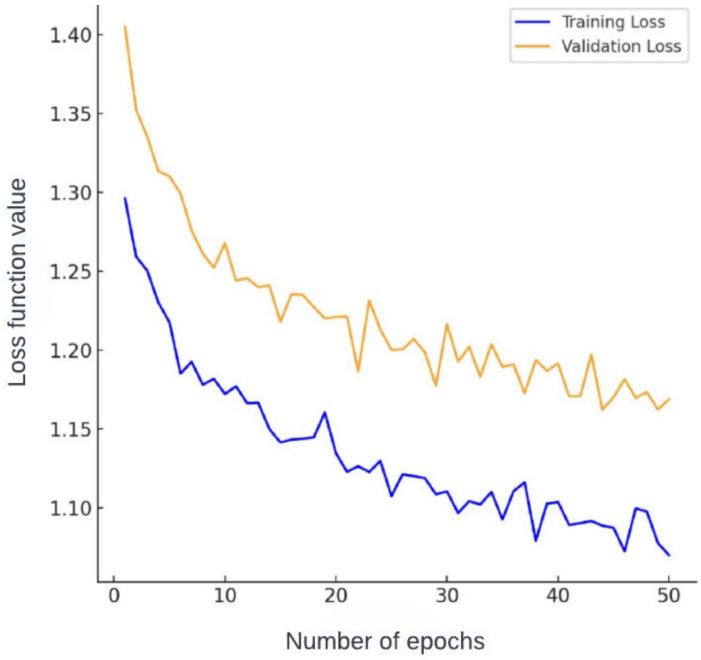


Figure 3.16: VGG16 model training and validation loss.

Table 3.4: Hyperparameters for proposed classification model and pre-trained models

S.No	Model	Total Parameters
1	Google Insception	**23,851,784**
2	VGG16	138,357,544
3	ResNet50	25,636,712

consistently sustaining strong recall while delivering mediocre performance in other domains. The aforementioned disparities underscore the need to carefully choose a model that aligns with the specific criteria in seed classification endeavors.

The table 3.4 presents the aggregate count of parameters for three distinct models, namely Google Inception, VGG16, and ResNet50. The hyperparameters of the models indicate their respective capacities. Specifically, Google Inception has the lowest capacity, with a total of 23,851,784 parameters. Following this, ResNet50 exhibits a slightly higher capacity, with 25,636,712 parameters. Lastly, VGG16 demonstrates the biggest capacity among the above models, with a total of 138,357,544 parameters. The diverse range of hyperparameter values demonstrates disparities in the complexity of the model, which in turn affects parameters such as the duration of training and the computer resources needed.

In summary, the chapter introduces novel way to identify and categorize seeds using Mask RCNN, which is critical for optimizing crop productivity and health. The suggested model employs seed pictures to leverage the enhanced capabilities of Mask RCNN, enabling for automatic temperature, moisture, humidity, and light modification depending on seed type. The model outperforms the state-of-the-art pre-trained models such as ResNet50, Google Inception, and VGG16 incorporated into the Mask RCNN architecture, with an accuracy of 86%. This demonstrates the efficacy of the suggested method for improving seed quality evaluation thereby contributing to the overall success of plant growth and development.

Chapter 4

SeedAI: A Novel Seed Germination Prediction System using Dual Stage Deep Learning Framework

This chapter introduced a novel two-stage network is proposed which leverages various convolutional Neural Networks (CNN) to automate detection of seeds and the assessment of their germination state. In the first stage Mask R-CNN framework is used for instantaneous segmentation of seeds and in the next stage this RoI is given as input to the proposed CNN model for germination prediction. This approach inherently extends the Mask-RCNN concept introduced in (Chapter 3), by using dual stage deep learning framework to enhance the overall performance.

4.1 Proposed Dual Stage Method for Predicting Seed Germination

In this section, a novel dual stage method is proposed for predicting the seed germination.

To improve the prediction of seed germination performance in this work a novel dual stage model is proposed. The proposed Detectron and Classification framework shown in

figure 4.1

1. A growth chamber system is designed specifically for collecting data set of seeds during their germination stage .

2. A novel two stage network is proposed to accurately identify seeds within petri dishes and classify them based on their germination status.

3. The proposed model is compared with other state-of-the-art models.

For effective automation of seed germination process a two stage network based on deep learning is proposed. The proposed model consists of Detectron2 in first stage for instantaneous segmentation of seed and novel CNN architecture for classification of seed germination.

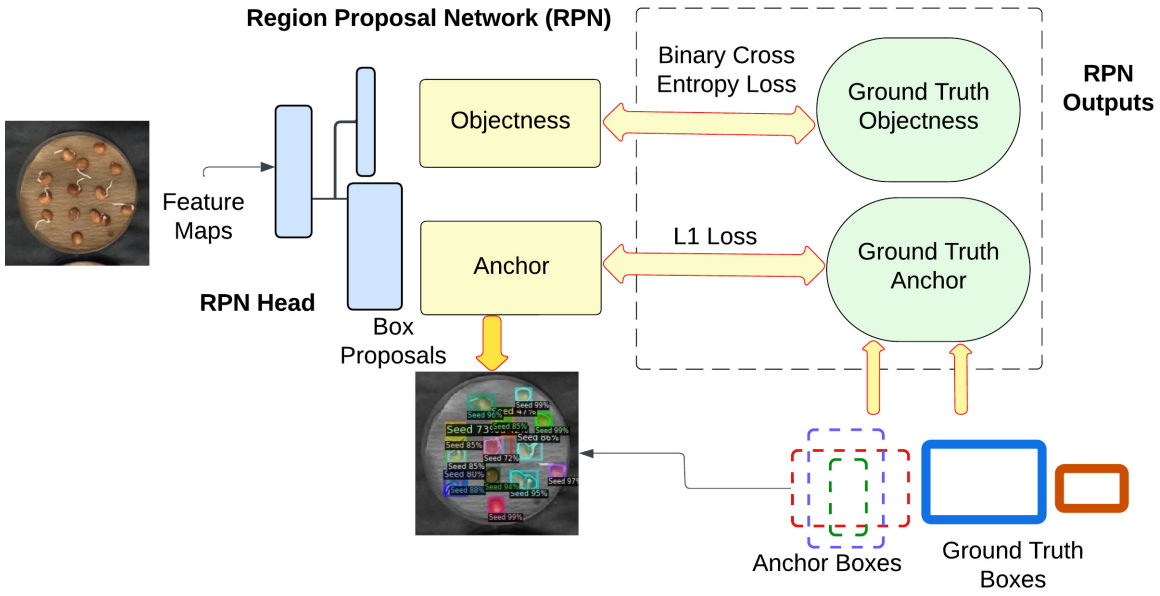


Figure 4.1: Block diagram of proposed two stage dual stage deep learning framework for seed germination prediction.

4.1.1 Seed Extraction using Detectron2

4.1.2 Module Selection

The first step involves selecting an appropriate anchor generator from the ANCHOR GENERATOR REGISTRY. This selection is crucial as it determines the base configuration for anchor generation.

4.1.3 Essential Parameters

Parameters that will influence the anchor generation process are:

1. Set of anchor sizes
2. Set of aspect ratios
3. Set of strides

These parameters are essential for configuring the anchor generator.

4.1.4 Configuration

After selecting the anchor generator module, the next step is to configure it with specific parameters. This configuration directly impacts how anchors are generated across the image.

The size and shape of each anchor can be defined using the following parameters:

- Sizes (A): A list of sizes (e.g., $[64^2, 128^2, 256^2]$) representing the area of the square anchors in pixels.

- Aspect Ratios (R): A list of aspect ratios (e.g., $[0.5, 1, 2]$) defining the width-to-height ratio of the anchors.

- Strides (S): The stride of each feature map level, determining the spatial resolution at which anchors are placed.

Given a feature map of width W and height H, and a stride S, the number of anchor positions (grid cells) on the feature map can be calculated as:

$$\text{Number of Positions} = \left(\frac{W}{S}\right) x \left(\frac{H}{S}\right) \tag{4.1}$$

For each position as shown in equation 4.1, anchors are generated based on the combinations of sizes and aspect ratios:

$$\text{Anchor Width} = \sqrt{A \times R} \tag{4.2}$$

$$\text{Anchor Height} = \frac{AnchorWidth}{R} \tag{4.3}$$

The specified anchor generator, configured with the desired parameters, is utilised within the RPN to generate anchors across the feature map. In seed Detection, customizing the anchor generator can significantly improve detection performance. Adjusting the sizes and aspect ratios to match the typical dimensions of seeds can lead to more accurate detection results.

4.2 RoI Refinement Techniques in Detectron2

Figure 4.2 illustrates the sequential refinement steps of an object detection pipeline applied to seed detection. The pipeline is based on the Detectron2 framework, which is a state-of-the-art object detection software system. The images represent a typical seeds in a petri dish and demonstrate the efficacy of the post-processing techniques at the inference stage of the model.

4.2.1 Preliminary Detection

The figure 4.2 (a) exhibits the initial detection phase, where the model has identified all possible ROIs corresponding to potential seeds. Each ROI is demarcated with a colored bounding box and an associated detection score. The plethora of overlapping boxes indicates an extensive raw detection output before the application of any filtering criteria.

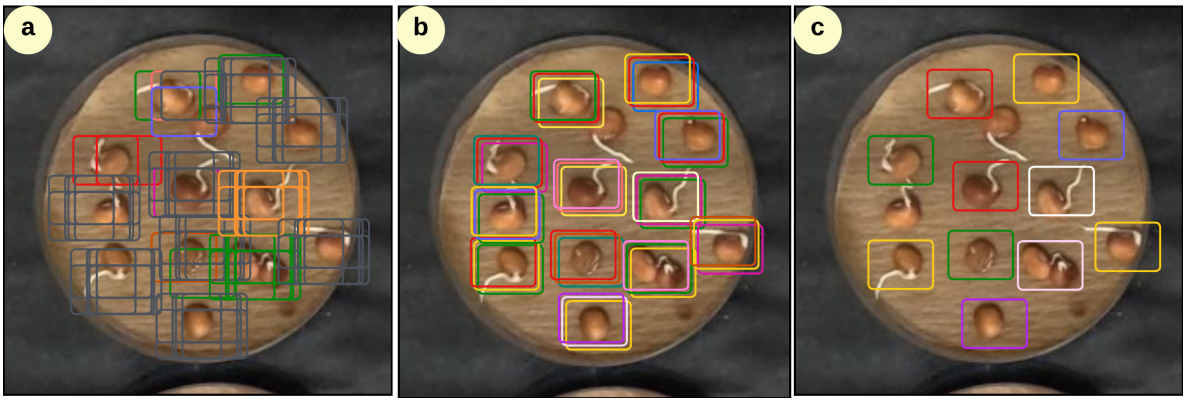


Figure 4.2: Efficacy of the post-processing technique (a). Intial detection phase (b). Reduced set of anchors (c). Single anchor box

4.2.2 Score Thresholding

In figure 4.2 (b), we observe a reduced set of bounding boxes. This reduction is the result of applying a score thresholding technique, which eliminates ROIs with confidence scores below a certain threshold. The threshold is determined empirically to balance the trade-off between recall and precision in seed detection. The remaining boxes indicate potential seeds with a higher likelihood based on the model's confidence.

4.2.3 Non-Maximum Suppression

Figure 4.2 (c) showcases the outcome after the NMS algorithm is applied. This post-processing step resolves the issue of multiple bounding boxes corresponding to the same seed by retaining only the box with the highest confidence score and suppressing the less confident, overlapping boxes. The result is a streamlined and accurate representation of seed detections, with minimal overlap, facilitating precise seed counting and localization.

The progressive refinement of the detection output, as depicted in these images, underscores the importance of post-processing in seed detection tasks. For applications such as seed detection in ecological monitoring, these steps are crucial for ensuring the accuracy and reliability of the results.

4.3 Loss Functions in Detectron2

To analyze the application of loss functions, specifically L1 Loss and binary cross-entropy loss, within the context of seed detection through machine learning models, mathematical underpinnings and practical implications of these approaches are explored.

4.3.1 L1 Loss Function

The L1 loss, often referred to as Least Absolute Deviations (LAD) or the Manhattan distance, quantifies the discrepancy between the predicted values ($\hat{y}_i$) by a model and the actual, observed values (y_i). It is mathematically represented as:

$$\text{L1 Loss Function} = \sum_{i=1}^{n} |y_{\text{true}} - y_{\text{predicted}}| \tag{4.4}$$

The equation 4.4 ensures that the loss is always a positive value, emphasizing the absolute difference without considering the direction of the error. In the context of seed detection, L1 Loss can effectively measure the accuracy of predicted seed locations or quantities against ground-truth data, penalizing any deviation irrespective of its nature.

4.3.2 Binary Cross-Entropy Loss

Binary Cross-Entropy Loss, also known as Log Loss, measures the performance of a classification model whose output is a probability value between 0 and 1. It is particularly suited for binary classification problems, such as distinguishing between seed and non-seed elements in an image. The loss function is shown in equation 4.5

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N} [y_i \log(\hat{p}_i) + (1 - y_i) \log(1 - \hat{p}_i)] \tag{4.5}$$

Here, y_i is the actual label (1 for seed, 0 for non-seed), $\hat{p}_i$ is the predicted probability of the i-th example being a seed, and N is the number of observations. This formula penalizes predictions that diverge from the actual labels, with the penalty increasing significantly for

predictions that are confidently wrong.

In the context of seed detection, binary cross-entropy is particularly helpful in distinguishing seeds (positive class) from the background (negative class) on a pixel or object level, depending on the task specifics (segmentation or detection).

The binary cross-entropy loss is a measure used for binary classification tasks. Mathematically, it is defined for a single example in equation 4.6

$$L(y, \hat{y}) = -\left(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})\right) \tag{4.6}$$

where - y is the true label (1 for seed, 0 for non-seed). - $\hat{y}$ is the predicted probability of the class being a seed. For a batch of N examples, the loss can be averaged as equation 4.7

$$L = -\frac{1}{N} \sum_{i=1}^{N} \left(y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)\right) \tag{4.7}$$

4.3.3 Binary Cross-Entropy Loss Function in Seed Detection

4.3.3.1 Model Output

Detectron2, when configured for seed detection, outputs predictions in terms of probabilities indicating the likelihood of seed presence at different locations or instances within an image.

4.3.3.2 Loss Calculation

For each prediction, the binary cross-entropy loss is calculated between the predicted probability and the ground truth label (seed or no seed). This process involves comparing the model's output against known labels in the training data.

4.3.3.3 Backpropagation and Optimization

The calculated loss is then used to adjust the model parameters through backpropagation. The goal is to minimize this loss, thereby improving the model's ability to distinguish between seeds and non-seed background.

4.3.3.4 Training Dynamics

Throughout the training process, the model learns to reduce the binary cross-entropy loss by improving its predictions. This involves learning features specific to seeds, such as shape, size, and texture, which are critical for accurate detection.

4.3.3.5 Evaluation and Adjustment

The model's performance is periodically evaluated on a validation set. Adjustments to the training process, such as learning rate changes or data augmentation strategies, may be made based on validation performance to further minimize loss and improve detection accuracy.

4.3.3.6 Loss Functions in Detecting Seeds

L1 Loss is directly applied to regression tasks where the goal is to predict specific attributes of seeds, such as size or count, from image data. It is straightforward and computationally efficient, making it suitable for models that need to run on limited hardware resources.

Binary Cross-Entropy Loss is utilized in classification tasks within seed detection, such as distinguishing seeds from the surrounding environment in a binary manner. This loss function is critical for training models to accurately classify pixels or regions of an image as containing seeds or not, optimizing the model's ability to predict with high confidence the correct class of each input as shown in equation 4.8.

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \tag{4.8}$$

4.3.3.7 Binary Cross-Entropy / Log Loss

The choice between L1 Loss and BCE loss depends on the specific requirements of the seed detection task at hand. For tasks requiring precise numerical predictions, L1 Loss is preferred due to its robustness to outliers and simplicity. Meanwhile, BCE loss is favored in binary classification tasks for its effectiveness in handling probabilities and facilitating the model's

ability to learn from imbalanced datasets, which is common in seed detection scenarios where the presence of seeds may be sparse or irregular.

The application of L1 Loss and BCE loss functions in seed detection leverages their unique advantages, whether it's the straightforward calculation and interpretation of L1 or the probabilistic nature of BCE, to enhance the accuracy and reliability of seed detection models. Through careful selection and implementation of these loss functions, researchers and practitioners can significantly improve the performance of machine learning algorithms in automated seed detection applications, contributing to advances in agricultural technology and research. The binary cross entropy loss is compared with predicted probability is shown in in figure 4.3. The model summary of Detectron model for seed and mask detection is shown in table 4.1.

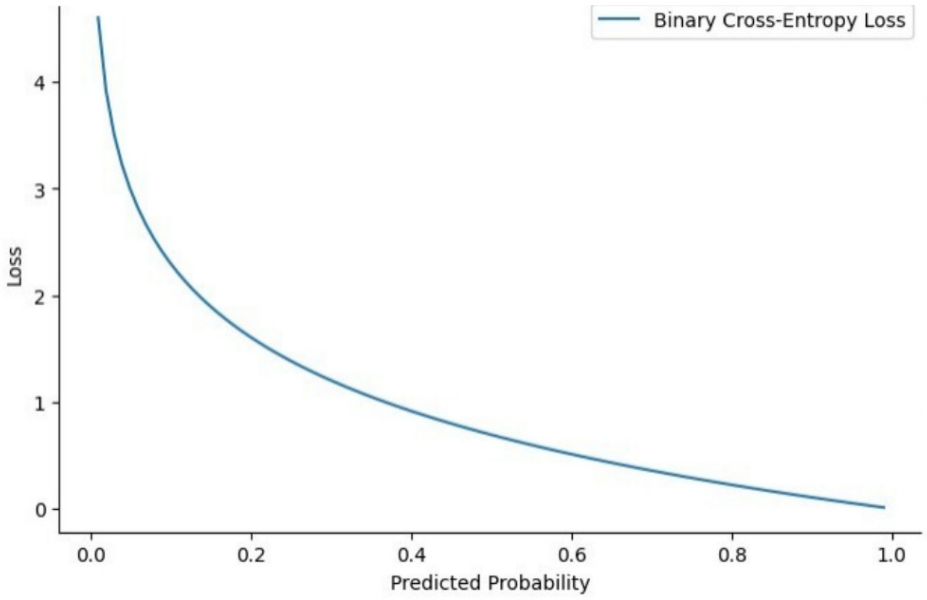


Figure 4.3: Binary cross-entropy loss vs. predicted probability

4.3.4 Seed Germination Classification using CNN

The second stage encompasses the proposed CNN model. The purpose of this model is to classify seeds based on their germination characteristics. Architecture of the proposed classification model is shown in table 4.2. It consists of a Conv2D layer comprising six

Table 4.1: Model summary of Detectron with layer sizes

Component	Layer Name	Input Size	Output Size
Backbone	FPN Lateral2	(256, 200, 200)	(256, 200, 200)
	FPN Output2	(256, 200, 200)	(256, 200, 200)
	FPN Lateral3	(512, 100, 100)	(256, 100, 100)
	FPN Output3	(256, 100, 100)	(256, 100, 100)
	FPN Lateral4	(1024, 50, 50)	(256, 50, 50)
	FPN Output4	(256, 50, 50)	(256, 50, 50)
	FPN Lateral5	(2048, 25, 25)	(256, 25, 25)
	FPN Output5	(256, 25, 25)	(256, 25, 25)
	Top Block	Varies	Varies
	Bottom Up ResNet	(3, 800, 800)	Varies
Region Proposal	RPN Head Conv	(256, 50, 50)	(256, 50, 50)
	RPN Head Objectness Logits	(256, 50, 50)	(3, 50, 50)
	RPN Head Anchor Deltas	(256, 50, 50)	(12, 50, 50)
ROI Heads	Box Pooler	Varies	(256, 7, 7)
	Box Head FC1	Varies	(1024,)
	Box Head FC2	(1024,)	(1024,)
	Box Predictor Cls Score	(1024,)	(2,)
	Box Predictor BBox Pred	(1024,)	(4,)
	Mask Pooler	Varies	Varies
	Mask Head FCN1	Varies	(256, 28, 28)
	Mask Head FCN2	(256, 28, 28)	(256, 28, 28)
	Mask Head FCN3	(256, 28, 28)	(256, 28, 28)
	Mask Head FCN4	(256, 28, 28)	(256, 28, 28)
	Mask Head Deconv	(256, 28, 28)	(256, 56, 56)
	Mask Head Predictor	(256, 56, 56)	(1, 56, 56)

filters of dimensions 3x3, which yields an output shape of (None, 28, 28, 6) and possesses a total of 896 parameters that can be adjusted during the training process. Subsequently, a MaxPooling2D layer is implemented, utilising a pool size of 2x2, thereby yielding an output shape of (None, 14, 14, 6). Next, a Conv2D layer with 16 filters, each having a size of 3x3 is appended. This layer produces an output shape of (None, 10, 10, 16) with 2416 toal trainable parameters. A subsequent MaxPooling2D layer is implemented with a pooling size of 2x2, which yields an output shape of (None, 5, 5, 16). Subsequently, the output tensor undergoes flattening through the utilisation of the flatten layer, leading to a resultant shape of (None, 120). Finally, a dense layer comprising of 128 units and the Rectified Linear Unit (ReLU) activation function is incorporated. This yielding an output shape of (None, 84) with 320002 trainable parameters.

Table 4.2: Architecture of proposed classification model.

Layer	**Output shape**	**Parameters**
Conv2d6 (Conv2D)	multiple	896
Maxpooling 2D6	multiple	0
conv2d7 (Conv2D)	multiple	18,496
Maxpooling 2D7 (maxpooling 2d)	multiple	0
flatten3 (Flatten)	multiple	0
dense	multiple	320,002
Total		**339,394**

Lastly, the model includes a Dense layer comprising of two units and utilises the softmax activation function to represent the output classes. The layer's output shape is defined as (None, 2), indicating that it produces a tensor with an unspecified number of rows and 2 columns. The proposed CNN model consists of a total of 3,39,394 parameters, all of which are shown in table 4.2. This CNN model is used for classifying seeds as either germinated or non-germinated.

Once the seeds in the petri dish have been detected using object detection with Detectron2, the subsequent procedure involves the classification of each seed as either germinated or non-germinated. The utilisation of a CNN model can facilitate the attainment of this objective.

The seed RoIs that have been identified are extracted from the image of the petri dish.

The RoIs encompass the individual seed images. The seed images undergo pre-processing by being resized to a predetermined size and having their pixel values normalised. This process guarantees that the data is standardised to a uniform structure that is appropriate for the CNN model. The seed images that have undergone pre-processing are subsequently inputted into the CNN model. The CNN is composed of several convolutional and pooling layers, which are designed to acquire and identify significant features from the initial images. The extracted features undergo processing in FC layers, which are responsible for performing the classification task. The layers facilitate the conversion of features into class probabilities, which serve as indicators of the probabilty of germination for each individual seed. During the training process, the CNN model is trained using labelled data. This labelled data consists of seed images, each of which is associated with a germination label indicating whether the seed has germinated or not.

The parameters of the model are optimised by employing loss functions, such as cross-entropy, with the objective of minimising the discrepancy between the predicted labels and the true labels. During the process of inference, the CNN model that has been trained is utilised to take the pre-processed seed images as its input. The model then proceeds to make predictions regarding the germination status of each individual seed. The classification label (germinated or not germinated) can be determined based on a predefined threshold using the class probabilities obtained from the output layer.

By employing a proposed CNN architecture and utilising a meticulously annotated dataset comprising images of both germinated and non-germinated seeds, it becomes feasible to effectively categorize the identified seeds within the petri dish, thereby furnishing significant insights pertaining to their germination status.

4.4 Experimental Results

4.4.1 Experimental Setup

The proposed system is a two-stage model utilised for the prediction of seed germination status.

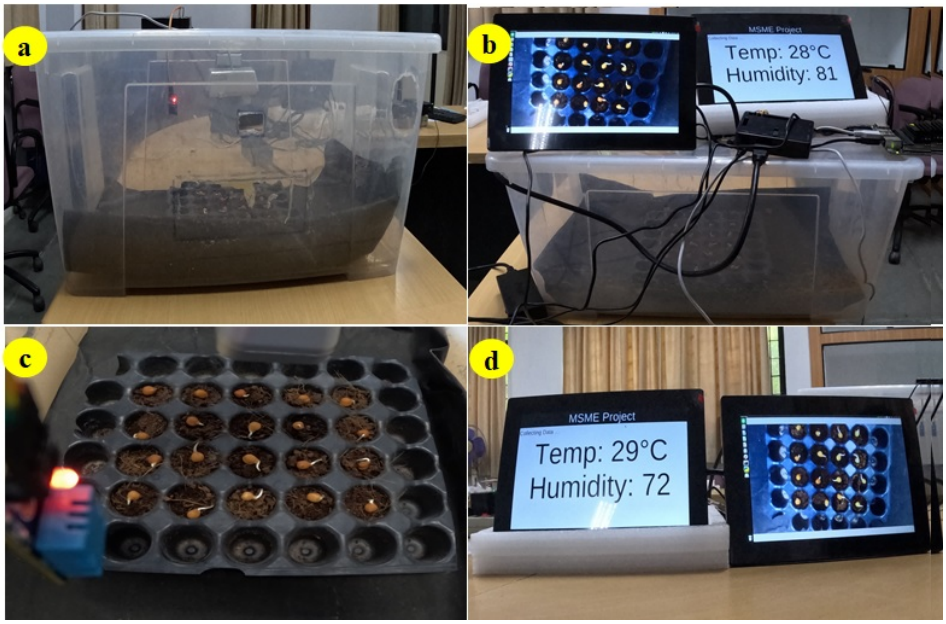


Figure 4.4: Prototype developed for seed germination prediction (a). Automated system for seeds data collection (b). Raspberry Pi [28] board collecting images and displaying temperature and humidity (c). Germinated seeds tray with temperature and humidity sensor (d). Maintaining constant temperature and humidity and display on touch screen.

4.4.2 Dataset Collection

A automated growth chamber is setup to collect the seed germination dataset which is shown in Figure 4.4. Figure 4.4 (a) provides an overview of the automated system set up for seamless data collection on seed germination rates. Figure 4.4 (b) depicts the Raspberry Pi board in action, capturing images for data analysis while simultaneously monitoring and displaying the current temperature and humidity levels, essential for maintaining an optimal growth environment. Figure 4.4 (c) shows a tray of germinated seeds where the temperature and

humidity sensors are strategically placed to ensure accurate environmental readings. Finally, Figure 4.4 (d) highlights the system's capability to maintain a constant temperature and humidity, crucial for seed germination, with real-time data displayed on a touch screen for easy monitoring and adjustments. The dataset presented in this article pertains to the germination of seeds in a growth chamber [73]. During which, the seeds underwent individual germination over a period of forty-eight hours. The experimental setup comprises a germination tray with 4 petri dishes, each containing more than ten seeds as shown in Figure 4.5 (a). Automated system is employed to regulate and sustain consistent levels of temperature, humidity, moisture, and light which indeed helps the seeds to germinate. Figure 4.5 (b) displays a sample image in our dataset obtained through an automated embedded system. The seeds are captured using a Raspberry pi [74] and camera and subsequently stored in database at regular intervals for 5 minutes. The procedure is iterated over a span of 3-4 days, depending upon the duration required for seed to be germinated. The collected seed data is annotated to provide a ground truth image for the purpose of training a deep learning model. The collected dataset distribution is mentioned in the table 4.3.

The dataset encompasses the germination process of red gram seeds through the utilization of an automated camera system. This system was programmed to record photos at regular intervals of 5 minutes. Thus, providing a granular view of the germination process. This repeated imaging displays a thorough path from initial swelling to radicle appearance and subsequent expansion. Quick intervals enable the early identification of germination and any anomalies, allowing for a fast response or analysis. Furthermore, the resulting rich dataset is invaluable for temporal research and improves the efficacy of deep learning models by providing many data points. Overall, such a thorough monitoring system provides critical information for both immediate observations and sophisticated analytical procedures.

4.4.3 Data Annotation

During annotation process a seed label is assigned to an image frame for the purpose of identifying the seed in a petri dish. It is done using robo flow tool [75]. This process is

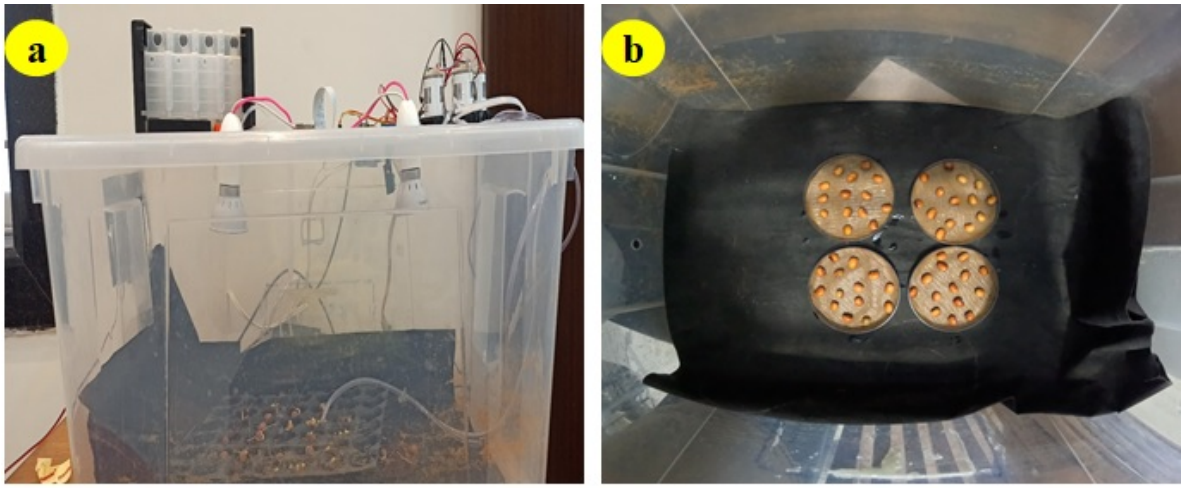


Figure 4.5: Seed germination prediction system (a) Model to collect images of seeds (b) Collected image by proposed model

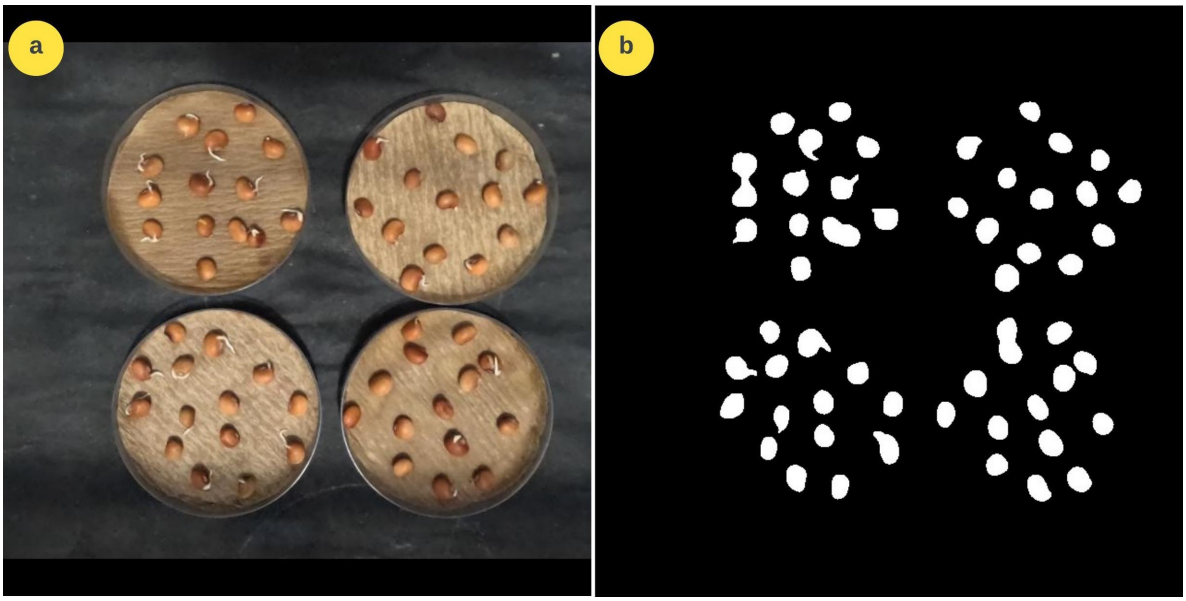


Figure 4.6: Example of generated seed dataset (a) seeds image (b) mask image tile of seed labels

illustrated in Figure 4.7. Figure 4.7 (a) shows an empty Petri dish, setting the baseline for our experiment. Figure 4.7 (b) presents the Petri dish filled with seeds, indicating the starting point for data collection. Figure 4.7 (c) depicts the seeds isolated from the Petri dish to emphasize the focus of our study. Figure 4.7 (d) demonstrates the meticulous annotation of individual seeds, a critical step for training our machine learning model to recognize and classify seed images accurately.

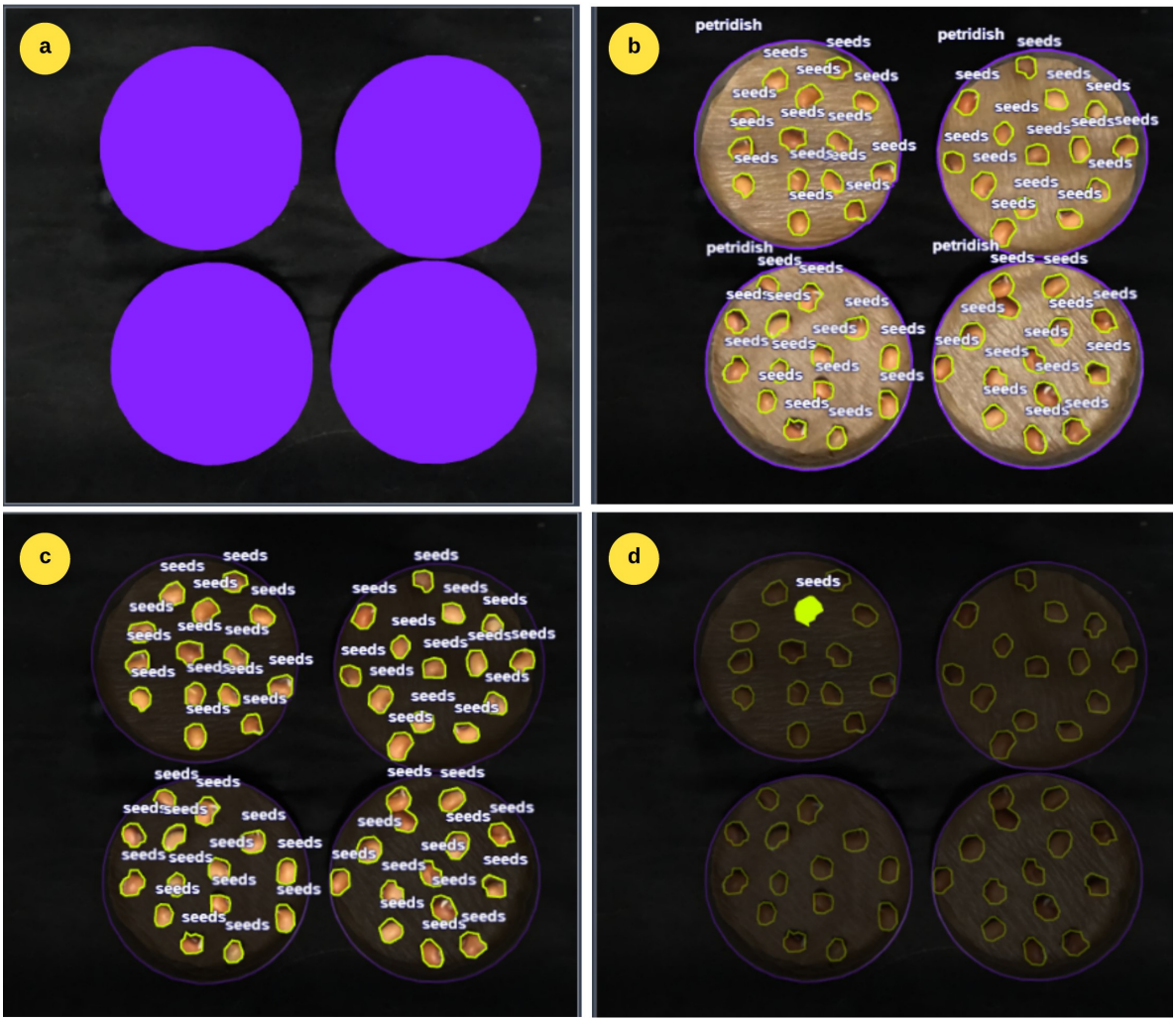


Figure 4.7: Annotating the seed dataset using robo flow tool. (a) Petri dish (b) Petri dish with seeds (c) Seeds without petri dish (d) Individual seed annotation.

The proposed model is evaluated using the Pytorch framework on a Tesla V100-SXM2 GPU with CUDA version 12.0. The Mask R-CNN is implemented the Detectron2 framework

Table 4.3: Dataset used for the proposed dual stage SeedAI model at different stages

S.No	# Images (Seeds)	Training	Testing	Total
1	Stage 1	572	147	719
2	Stage 2	858 (Germination), 858 (No Germination)	240	1955

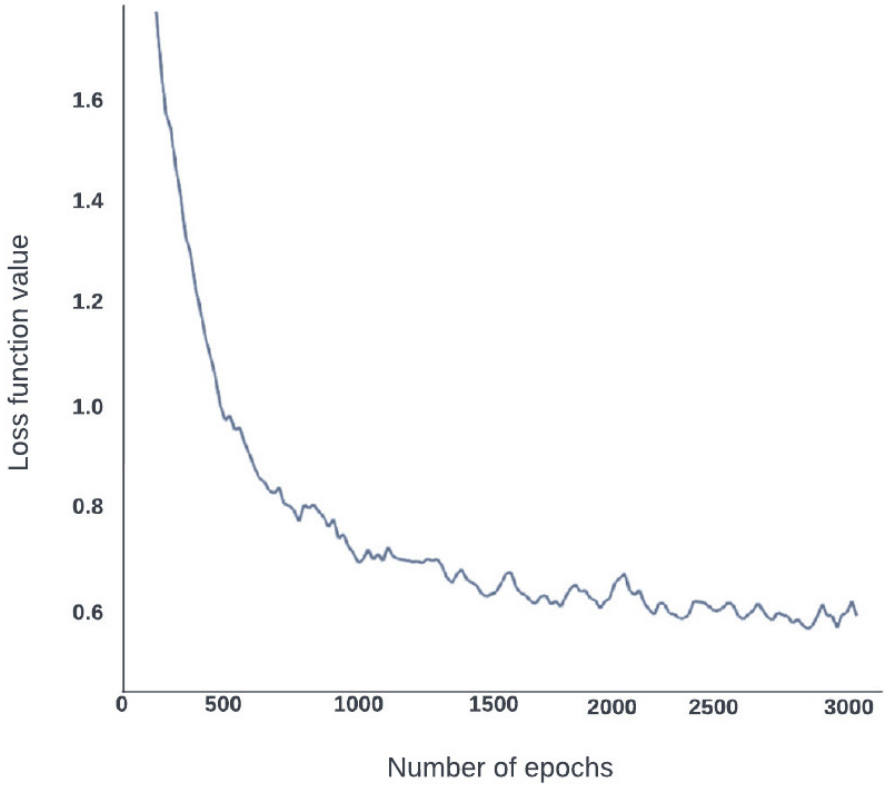


Figure 4.8: Proposed detectron model learning curve

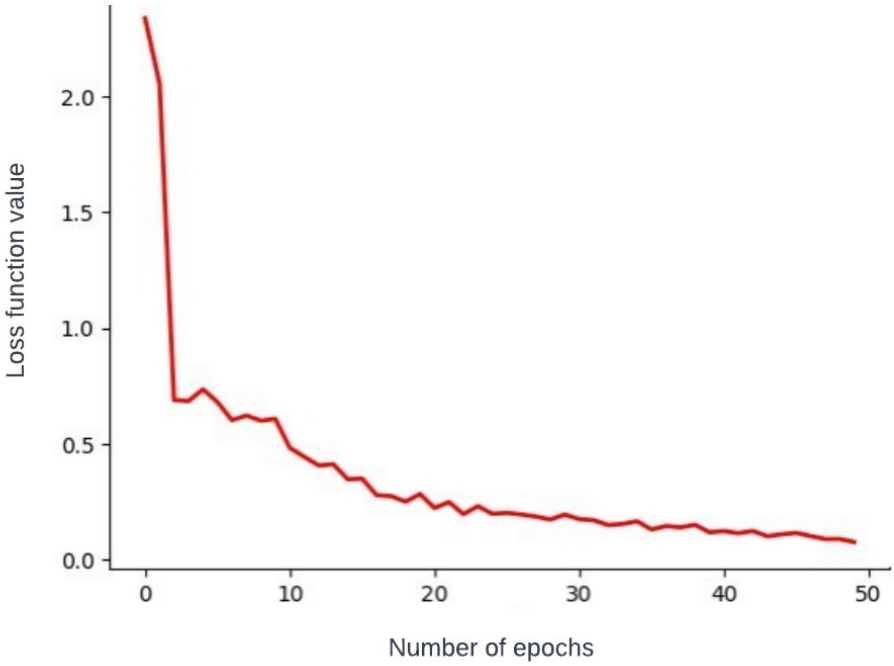


Figure 4.9: Seed classification learning curve

with a suitable backbone network architecture and conducted on the collected dataset.

4.4.4 Performance Measure

To assess the effectiveness of our proposed two stage network in predicting seed germination six widely used metrics, namely pixel accuracy, Intersection of Union (IoU), precision, recall, and F1 score are employed. The pixel accuracy refers to the proportion of pixels in the image that have been accurately segmented. The term "IoU" denotes the intersection between predicted masks and ground truth masks. Precision is the ratio of successfully extracted masks to all anticipated masks, whereas recall is the ratio of correctly recognised masks to ground truth. Finally, the F1 score is defined as the harmonic mean of precision and recall. These measures are already explained in chapter 3.

Proposed model demonstrates a pixel accuracy of 0.88, denoting that 88 % of the pixels in the predicted masks correspond with the ground truth masks. When compared with state of art models. ResNet50 demonstrates a pixel accuracy of 0.84, whereas LeNet exhibits a

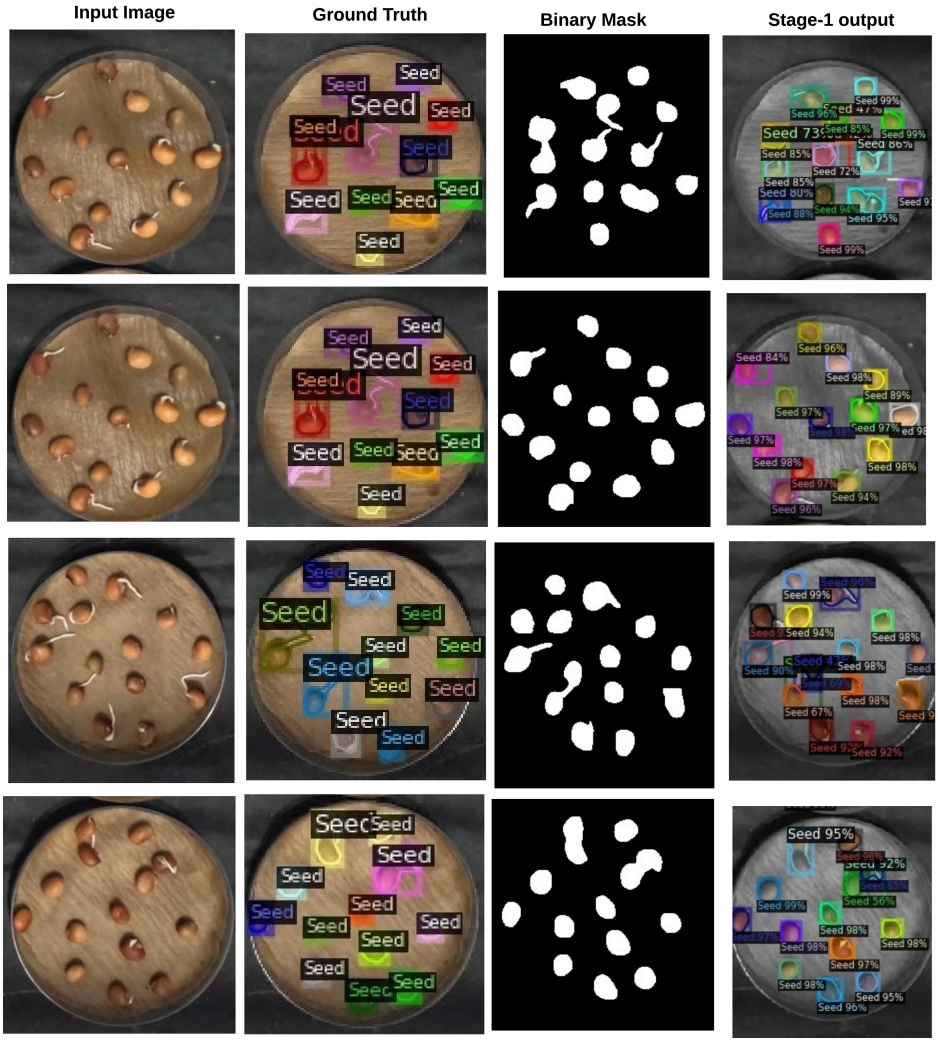


Figure 4.10: Models for detecting masks (illustration in a segmentation of instances)

Table 4.4: Evaluation metrics of the model with different pre-trained models

Model	Pixel Accuracy	IoU	Precision	Recall	F1
Proposed Model	**0.88**	**0.77**	**0.89**	0.85	0.87
ResNet50	0.84	0.75	0.79	**0.94**	0.86
Inception	0.85	0.45	0.49	0.85	**0.88**
LeNet	0.82	0.70	0.79	0.94	0.82

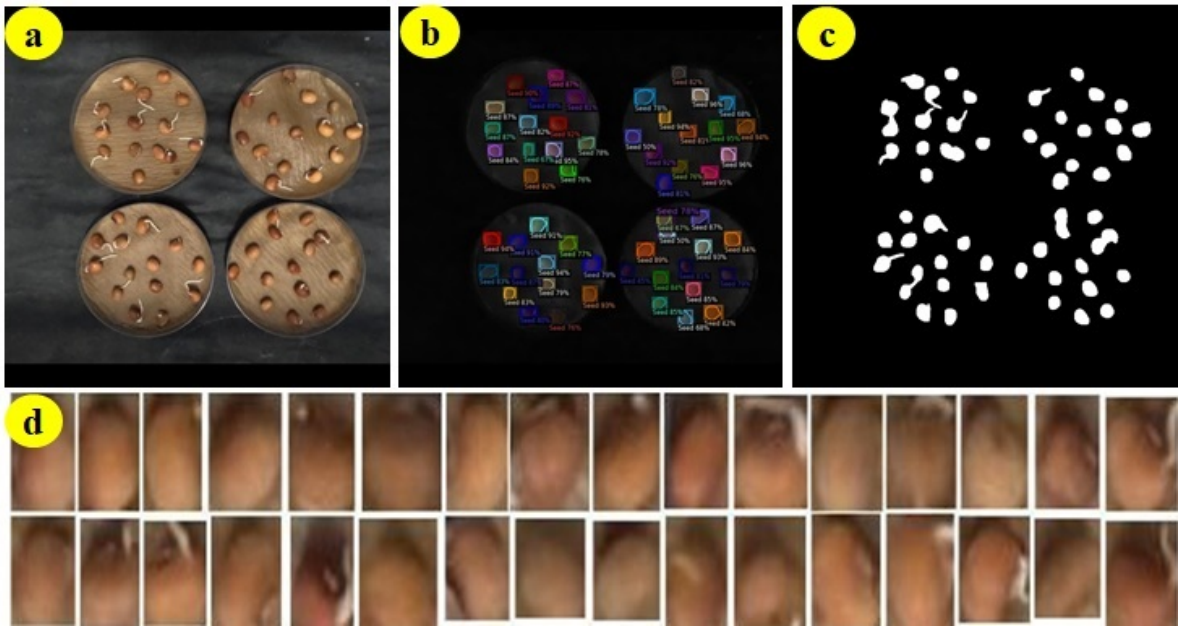


Figure 4.11: Input and output of the proposed SeedAI model (a). Input seed image for the SeedAI model. (b). Mask image generated using the SeedAI model (c). A binary mask is generated from the SeedAI model. (d). Segmented seed images from the petri dish

pixel accuracy of 0.82 and Google Inception pixel accuracy is 0.85. The Intersection over Union (IoU) metric reveals that proposed model exhibits an IoU value of 0.77, signifying a substantial degree of overlap between the predicted masks and the ground truth masks. ResNet50 demonstrates an Intersection over Union (IoU) value of 0.75, whereas LeNet exhibits an IoU value of 0.70 and Inception model has IoU value of 0.45 The precision score of proposed model is 0.89., signifying that 89% of the predicted positive instances are accurately classified. ResNet50 demonstrates a precision score of 0.79, whereas LeNet exhibits a precision score of 0.79 and google inception is of 0.49. The recall score of proposed system is 0.85, indicating that it correctly identifies 85% of the actual positive instances. ResNet50 demonstrates a recall score of 0.94, whereas LeNet exhibits a recall score of 0.94 and google inception is 0.85. Proposed system exhibits an F1 score of 0.87, which represents the harmonic mean of its precision and recall metrics. ResNet50 demonstrates an F1 score of 0.86, whereas LeNet exhibits an F1 score of 0.82 and google inception is 0.82. The comparision is done in table 4.4

Table 4.5: Hyper parameters for proposed model and pre-trained models

S.No	Model	Total Parameters
1	Proposed Model	**3,39,394**
2	LeNet	61,326
3	ResNet50	24,637,826
4	Google Inception	23,903,010

4.4.5 Result Analysis

Further, the proposed model is compared in terms of model complexity with state-of-the-art models. These results are furnished in table 4.5 and best values are highlighted. From the results it is observed that proposed model achieved on-par performance 88% with Google inception by reducing model complexity drastically by 98.5% which is significant move.

4.4.6 ROC Curve

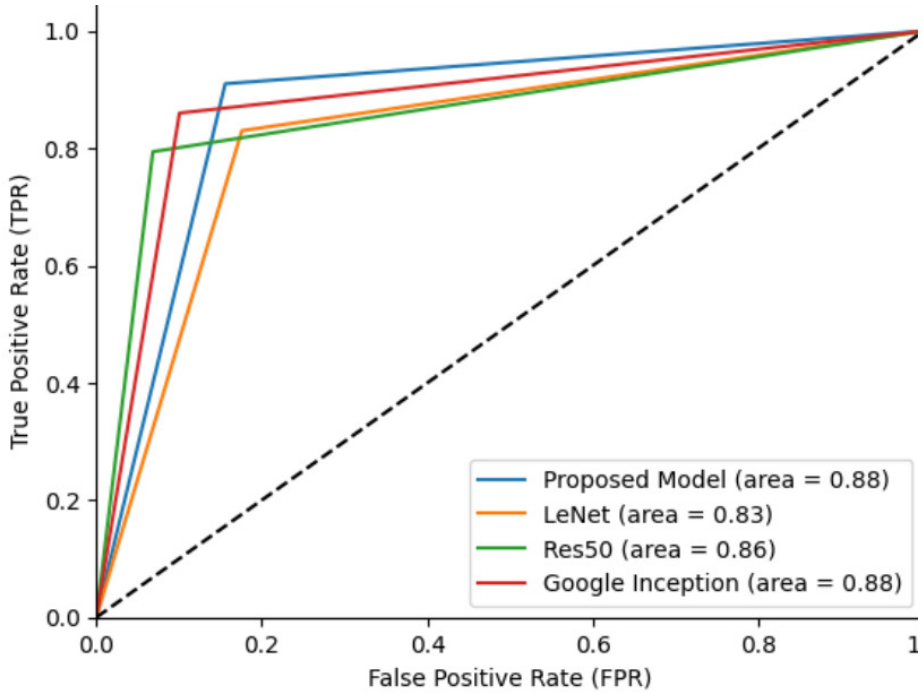


Figure 4.12: ROC to evaluate the performance of different models

The Receiver Operating Characteristic (ROC) curve is commonly employed in the as-

sessment of the efficacy of binary categorization models. This diagram depicts the correlation between the True Positive Rate (TPR) and the False Positive Rate (FPR) across different classification thresholds. To compare proposed model with ResNet50, LeNet and InceptionNet ROC curves are plotted graphically and shown in Figure 4.12. From the figure it is observed that our proposed model balances TPR and FPR when compared to other state-of-the-art models.

4.5 Summary

In agricultural domain seed germination is one of the main steps to assess the seed quality. It is used by seed industries, seed breeders and agricultural research institutes for production of high-quality seeds. However, the current research uses image processing techniques for seed germination and are mostly manual and labor intensive. Hence, a automatic monitoring system is essential to speed up the testing process. A novel approach is proposed for the automatic seed germination classification. The proposed system uses detectron2 in first stage for seed segmentation and CNN in second stage for classification.

The proposed novel two-stage network leverages various CNNs to automate detection of seeds and the assessment of their germination state. In the first stage detectron2 framework is used for instantaneous segmentation of seeds and in the next stage this RoI is given as input to the proposed CNN model for germination prediction.

The proposed model is trained and tested against own collected dataset. The proposed model got best performance in term of accuracy, IoU, Precision when compared with ResNet50, Inception Net and LeNet models. Moreover, the proposed model achieved on-par performance 0.88 with Google inception by reducing model complexity drastically by 0.98 which is significant move.

Chapter 5

Real-time Seed Detection and Germination Analysis in Precision Agriculture: A Fusion Model with U-Net and CNN on Jetson Nano

This chapter introduce a novel fusion model for seed detection and germination is proposed. The proposed model combines the U-Net and CNN architectures for seed segmentation and classification respectively. By harnessing U-Net's capabilities in image segmentation and CNN's strengths in classification, the proposed approach enables effective seed germination analysis. Additionally, the model is specifically optimized for real-time processing and applications by implementing it on the Nvidia Jetson Nano embedded GPU platform. This approach inherently extends the Mask R-CNN concept introduced in (Chapter 3), by using semantic segmentation to enhance the overall performance.

This chapter proposes a fusion model ported on jetson nano for predicting seed germination while fusion model shows promising results for seed germination by increasing the accuracy. The dual stage network proposed in chapter 4 marks an important step towards addressing this challenge. However, the fusion model optimized the process involves further

intricacies.

To extend the groundwork laid in chapter 4, the fusion model is introduced with semantic segmentation and proposed CNN. It aims to optimize a novel fitness function that considers both mean accuracy and mean diversity. By doing so, it not only address the challenge of base classifier selection as in chapter 4 but also introduce a mechanism for evaluating the performance of different ensemble combinations.

Chapter Organization: The proposed methodology is presented in Section 5.1. The experimental results are provided in Section 5.2. A summary of the work is described in Section 5.3.

5.1 Proposed Methodology

This work integrates U-Net [76], [77] and CNN architectures [78] to form a novel fusion model for seed detection and germination classification. Leveraging the strengths of U-Net for image segmentation and CNN for classification, this fused approach achieves practical seed germination analysis. The proposed hybrid model combines the U-Net architecture for seed image segmentation with CNN for germination state and abnormality classification. This combination technique offers improved accuracy, particularly in complicated datasets with varying seed shapes and unclear germination phases. The model's modular design, which separates segmentation and classification, enables independent fine-tuning, optimising accuracy, and successfully tackling unique difficulties.

Furthermore, the model is implemented on the Nvidia Jetson Nano embedded GPU platform to facilitate real-time processing and applications as shown in figure 5.1. This showcases the potential for practical deployment in real-world scenarios, such as monitoring in intelligent agriculture systems with masks of seeds at different germination states.

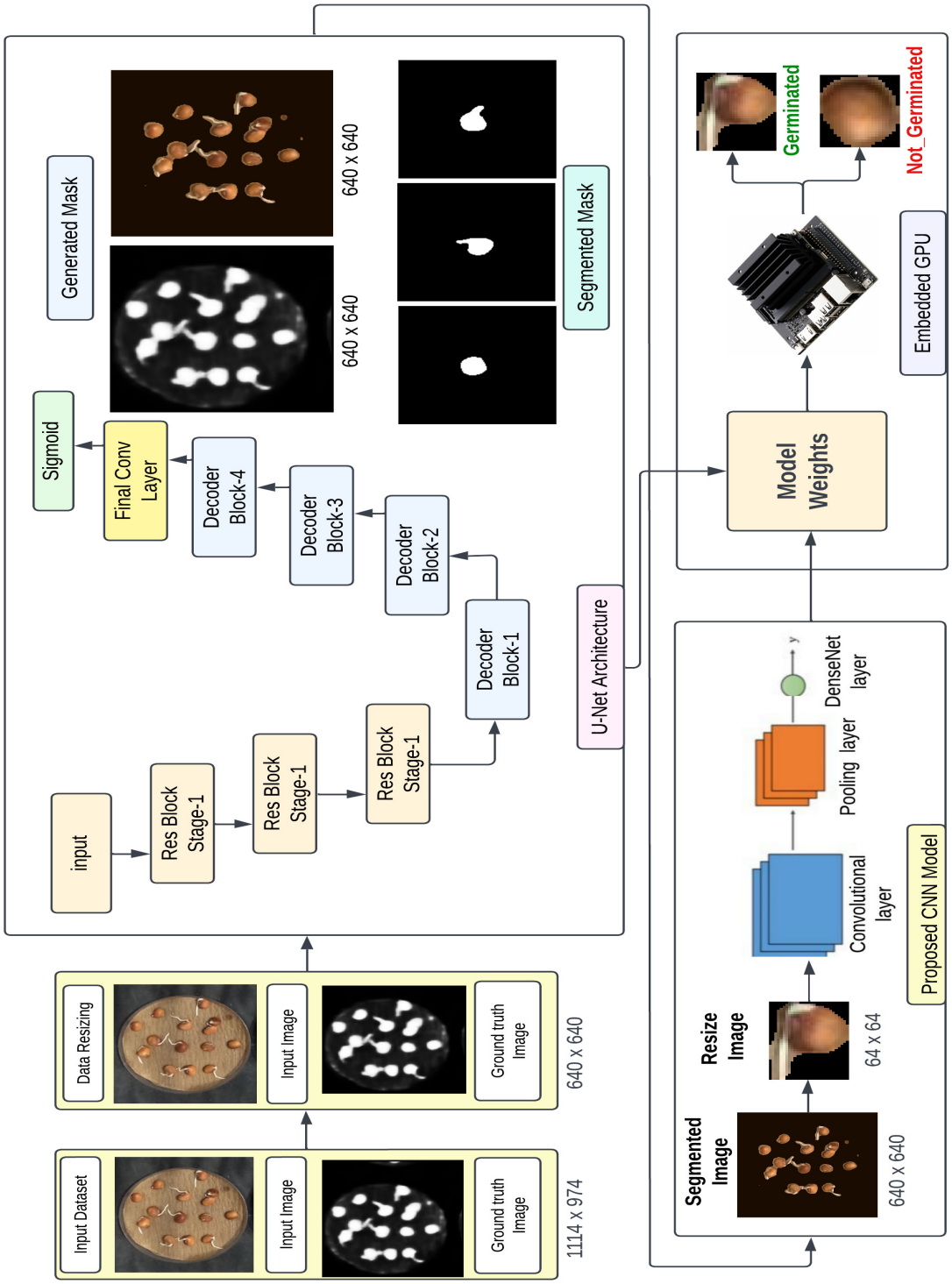


Figure 5.1: Seed germination classification on segmented images using U-net architecture

5.1.1 Semantic Segmentation of Seed using U-Net

U-Net CNN identify image pixels for semantic segmentation tasks. Its construction resembles a "U," thus the name U-Net. The architecture of U-Net is shown in table 5.1. The U-Net model would input a seed picture and produce a "mask" labeling each pixel as a seed. The contraction (encoder) and expansion (decoder) paths make up the U-Net design. Encoder path is a Convolution-ReLU stack followed by Max Pooling operation capturing picture context. A filter or kernel K of a given size slides across the input picture i.e seed image and conducts element-wise multiplication during convolution. The mathematical representation of encoder is given in equation 1.

$$Y[i,j] = X[i+m, j+n] * K[m,n] \tag{5.1}$$

where i, j are spatial dimensions and m, n kernel dimensions. ReLU activation followed by convolution. Max pooling reduces spatial dimensionality and controls over-fitting. Up-convolution, skip-connection, and Convolution-ReLU operations comprise the decoder. The transposed convolution procedure increases feature map spatial dimensions, helping the network learn to localize. The mathematical equation of decoder is given in equation 2.

$$Y[i,j,d] = X'[i,j,d] \oplus K[m,n] \tag{5.2}$$

where X and X' are the decoder and encoder feature maps, and 'd' is the feature map depth. U-Net's last layer is a 1X1 convolution followed by a softmax activation function to segment the seed image.

Seed segmentation is learned using images of seeds with each pixel labeled as "seed". The model recognizes seeds in image and applies this information. A binary mask would highlight the image's segmented seeds as shown in figure 5.2.

Table 5.1: Architecture of U-Net model.

Layer (Type)	Output Shape	Parameters
input2(Input layer)	None, 640, 640, 3	0
conv2d(Conv2D)	None, 640, 640, 16	448
conv2d1(Conv2D)	None, 320, 320, 16	2320
maxpooling2d(MaxPooling2d)	None, 320, 320, 16	0
conv2d2(Conv2D)	None, 320, 320, 32	4640
conv2d3(Conv2D)	None, 320, 320, 32	9248
maxpooling2d1(MaxPooling2d)	None, 160, 160, 32	0
conv2d4(Conv2D)	None, 160, 160, 64	18496
conv2d5(Conv2D)	None, 160, 160, 64	36928
max pooling 2d2 (MaxPooling2d)	None,80, 80, 64	0
conv2d6(Conv2D)	None,80, 80, 128	73856
conv2d7(Conv2D)	None,80, 80, 128	147584
maxpooling2d3(MaxPooling2d)	None,40, 40, 128	0
conv2d8(Conv2D)	None,40, 40, 256	295168
conv2d9(Conv2D)	None,40, 40, 256	590080
upsampling2d (Upsamling2D)	None,80, 80, 256	0
concatenate	None,80, 80, 256	0
conv2d10 (Conv2D)	None,80, 80, 128	442496
conv2d11(Conv2D)	None,80, 80, 128	147584
upsampling2d1(Upsamling2D)	None,160, 160, 128	0
concatenate1(Concatenate)	None,160, 160, 192	0
conv2d12(Conv2D)	None,160, 160, 64	110656
conv2d13(Conv2D)	None,160, 160, 64	36928
upsampling2d(Upsamling2D)	None,320, 320, 64	0
concatenate2(Concatenate)	None,320, 320, 96	0
conv2d14(Conv2D)	None,320, 320, 32	27680
conv2d15(Conv2D)	None,320, 320, 32	9248
upsampling2d3(Upsamling2D)	None, 640, 640, 32	0
concatenate3(Concatenate)	None, 640, 640, 48	0
conv2d16(Conv2D)	None, 640, 640, 16	6928
conv2d17(Conv2D)	None, 640, 640, 16	2320
conv2d18(Conv2D)	None, 640, 640, 1	17

5.1.2 Seed Germination Classification using Proposed CNN Model

CNN's can automatically and adaptively learn spatial hierarchies of characteristics from pictures, making them robust image classifiers. CNNs can classify seed pictures into germination phases. Images of seed germination would feed to the CNN in input layer. CNN's

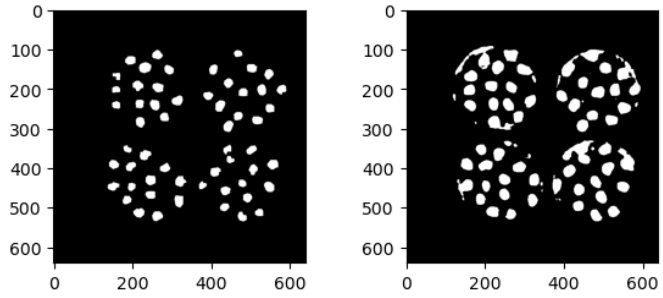


Figure 5.2: Ground truth image and segmented masks images of the U-Net model

Table 5.2: Architecture of the proposed CNN Model

Layer (Type)	**Output Shape**	**Param**
Conv2d6(Conv2D)	multiple	896
Maxpooling 2D6	multiple	0
conv2d7(Conv2D)	multiple	18496
Maxpooling 2D7(maxpooling 2d)	multiple	0
flatten3 (Flatten)	multiple	0
dense	multiple	320002
	Total	**3,39,394**

initial layers are convolutional. These layers "slide" filters or kernels across input pictures and execute element-wise multiplication to build a "feature map" or "convolved feature". This method extracts picture elements, including edges, corners, and textures, that may indicate seed germination in Convolutional Layers. The network decreases the spatial size of the feature map via a pooling operation like max pooling to control overfitting and minimize computational complexity in pooling layer.

Numerous rounds of convolution and pooling is performed, Then the CNN contains fully linked layers like a neural network. These layers enable the network to learn non-linear combinations of high-level features from CNN convolutional output. The last fully connected layer produces germination class probability distributions in fully linked layers. CNN classifies seed germination at output layer. The architecture of proposed model is shown in table 5.2.

Table 5.3: Redgram dataset used for the proposed fusion model

Training	Testing	Total
29,996 Seeds	5,875 seeds	35,871

5.2 Experimental Results

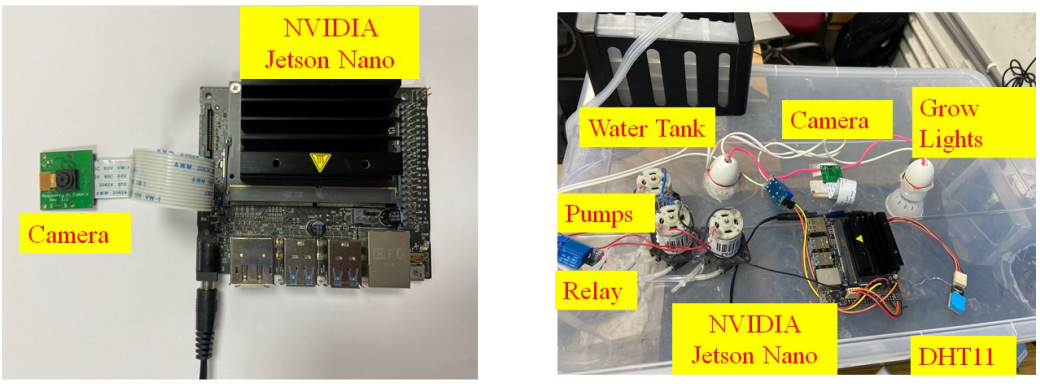
The proposed model is a fusion of U-Net and CNN for automated seed germination analysis, which runs on NVIDIA Jetson Nano embedded GPU as shown in figure 5.3.

5.2.1 Dataset Collection

This study utilised a combination of a Jetson Nano [79] and a camera interface [80] to collect the dataset. The Jetson Nano is a Artificial Intelligence (AI) computer that enables the efficient processing of images directly at the edge [81], thereby mitigating the necessity for data transmission and guaranteeing the acquisition of real-time data. The camera, which is connected to the Jetson Nano as shown in figure 5.3a, can capture high-resolution images depicting red gram seeds at different stages of germination as shown in figure. The aforementioned configuration facilitates uninterrupted surveillance and data acquisition throughout the germination process.

There are around 630 images in the dataset, and each contains 56 seeds. This creates 35,871 red gram individual seed images. We constituted 80% of images for training and 20% for test dataset. Then, this dataset is divided into 28,696 seed images are used for training, and 7,174 seed images are used as test data sets as shown in table 5.3.

The utilization of this setup for the collection and processing of data in real-time offers significant contributions to the understanding of the seeds germination process as shown in figure 5.3b. Consequently, it improves the precision and reliability of our model, thereby expanding its potential applications in the fields of precision agriculture.



(a) (b)

Figure 5.3: Prototype of the proposed fusion model (a) Camera interfaced to Jetson nano development board, (b) Growth chamber used for collecting data set.

5.2.2 Data Labelling

The proposed model 'ground truth images' pertains to precisely annotated images that accurately indicate the specific positions of seeds and their corresponding stages of germination. The images that are appropriately labelled serve as the standard or reference against which the model's predictions are evaluated during the training process.

To produce the ground truth images, it is necessary to annotate each image in the dataset manually. To facilitate seed detection, one approach could involve delineating each seed by employing bounding boxes or generating binary masks to differentiate seed regions within the image. In the process of categorising germination, each seed is assigned a distinct label corresponding to its current stage of germination as shown in figure 5.4. Figure 5.4 (a) of the figure showcases seeds at various stages of germination within a Petri dish, illustrating the diversity of germination progress captured in our dataset. Figure 5.4 (b) and (c) display the ground truth masks for these seeds, which are crucial for accurately labeling each seed according to its specific stage of germination. These ground truth masks serve as a reference for training our model.

The utilisation of ground truth images is a crucial component in the training procedure of the model. The U-Net component of the model acquires the ability to distinguish seeds from the background by comparing its segmentation output with the ground truth. In a similar

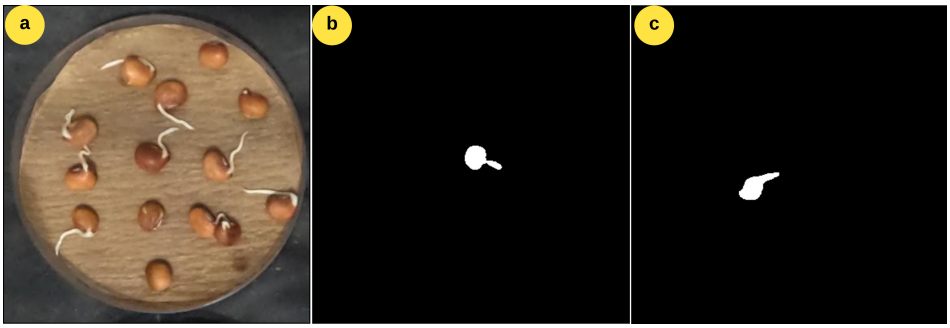


Figure 5.4: The input image and ground truth image of the proposed fusion model (a) Seed germination in petri dish (b)&(C) Ground truth masks of seeds.

manner, the CNN acquires the ability to classify the various stages of germination by evaluating its output in comparison to the accurate labels provided. Furthermore, the utilisation of these ground truth images is crucial in assessing the efficacy of your model. The performance metrics of the model, including accuracy, precision, and recall, can be assessed by comparing its predictions on new images with the ground truth. This evaluation involves combining the model's predictions against the actual values.

It is imperative to bear in mind that the generation of precise ground truth images can be a laborious process and may necessitate specialised expertise. However, it constitutes a crucial stage in the development of an efficient model.

5.2.3 Experimental Setup

The experimental setup involves the design and construction of a specialised growth chamber, taking dimensions, illumination, temperatures, and control of humidity. NVIDIA Jetson nano board is interfaced with a camera, temperature and humidity sensors with a display. Temperature and humidity controlled to maintain germination states for the seed. The retained moisture and temperature are depicted in figure 5.5 (b). The image acquired from the growth chamber is illustrated in the reference figure 5.5 (c). The overall system is shown in the figure 5.5 (a).

The Jetson Nano is connected to the MIPI CSI camera module. The imaging device can be accessed and managed via Python programming language, typically employing OpenCV. A

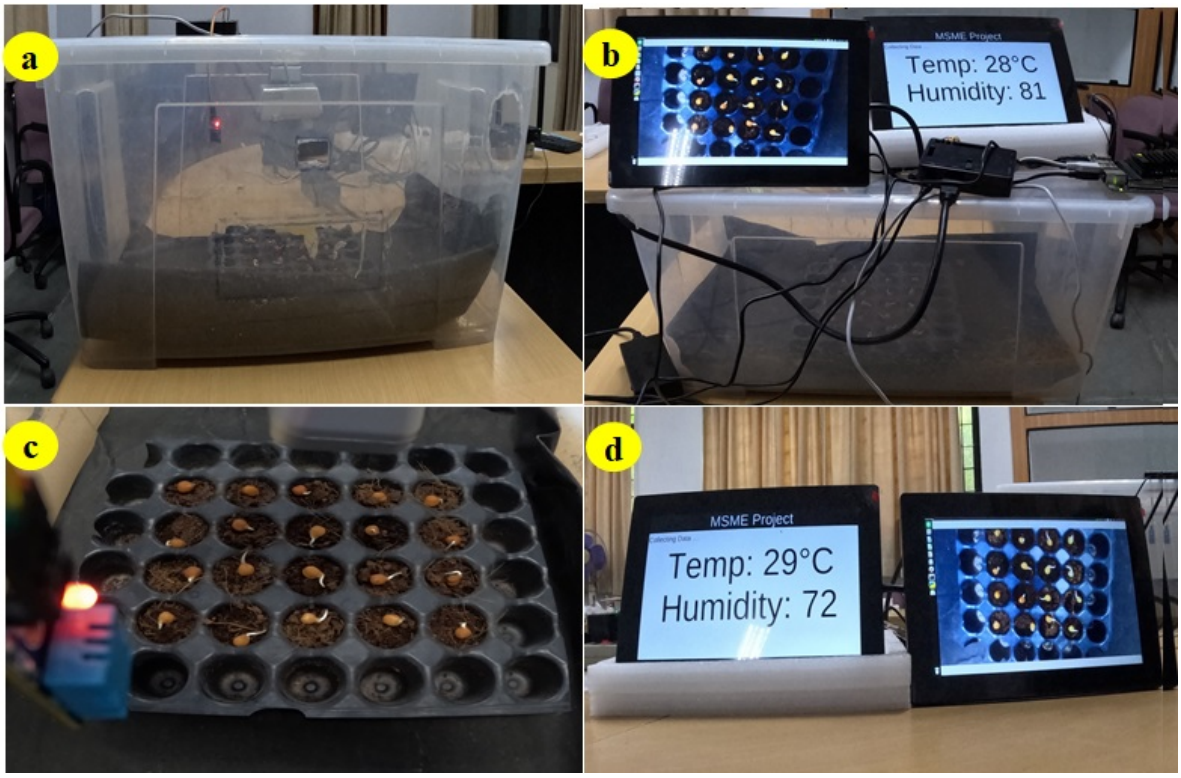


Figure 5.5: (a) Proposed growth chamber, (b) Temperature and humidity monitoring, (c) Collected seed images, and (d) Display of the proposed growth chamber

Python script is developed to facilitate the process of image capturing for every five seconds. This script continuously captures high-resolution images of the seeds as shown in in figure 5.5 (d). For classification and simple retrieval, each image is timestamped and stored as shown in figure 5.5 (d).

The acquired seed 35,871 images are subsequently subjected to pre-processing techniques and inputted into our fused U-Net and CNN model for the purpose of seed detection and germination classification. The compiled images provide an informative timeline of the process of seed germination. The image patterns can then be used to train and validate proposed fusion model that predicts the germination process based on the image patterns. By collecting images for every five minutes, the model receives granular, comprehensive data that enables it to detect and learn even subtle changes during germination and increase the accuracy of its predictions. To record the entire growth process of seed germination.

The collected dataset is used for fusion model employing U-Net and CNN for seed detec-

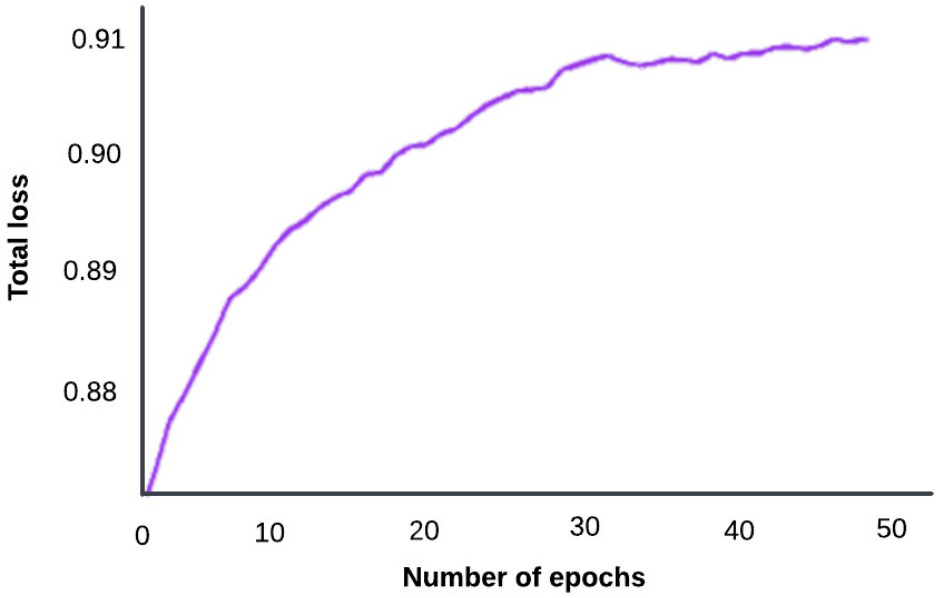


Figure 5.6: Learning curve for U-Net Model.

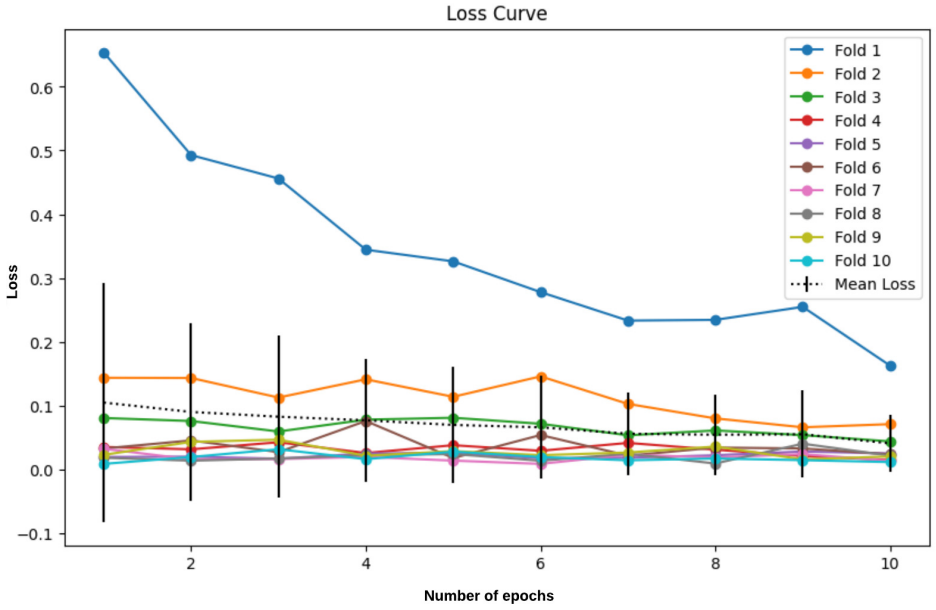


Figure 5.7: Fold wise loss curve for proposed CNN model

Table 5.4: 10 fold validation for proposed fusion model

Model	Pixel Accuracy	IoU	Precision	Recall	F1-Score
Fold-1	0.90	0.53	0.56	0.93	0.70
Fold-2	0.64	0.31	0.86	0.33	0.48
Fold-3	0.65	0.52	0.63	0.76	0.69
Fold-4	0.60	0.54	0.55	0.96	0.70
Fold-5	0.76	0.63	0.74	0.81	0.77
Fold-6	0.69	0.39	0.96	0.40	0.56
Fold-7	0.69	0.61	0.62	0.98	0.76
Fold-8	0.81	0.65	0.91	0.70	0.79
Fold-9	0.74	0.66	0.65	1.00	0.79
Fold-10	0.87	0.77	0.89	0.85	0.87
Avg	**0.91**	**0.84**	**0.90**	**0.92**	**0.91**

tion and classification. Using the prepared data set, the model is trained and then optimised with an appropriate loss function, optimizer, and performance metrics. The performance of the model is rigorously evaluated against the validation and test datasets, gauging its effectiveness and deployment suitability.

U-Net model is evaluated using Keras, Tensor Flow 2.12.0 on an HP Z6 G4 Workstation with 52 cores and 64 GB of DDR4 RAM. A U-Net model is trained using 35,871 seed images for 50 epochs on a GPU for semantic segmentation, the proposed CNN model is trained for 50 epochs on the GPU using germination and non-germination seed images.

The proposed system is trained on a system having specification, CPU as Intel Xeon GOLD 6226R @ 2.9 GHz, GPU as NVIDIA RTX A4000 with 64 GB internal memory.The proposed fusion model took roughly 6 hours for 50 epochs, a 1.3 MB size of model weight file is generated. The model weights are deployed on Jetson Nano, and it takes 35 seconds to process the input image for germination prediction.

To support real-time applications, the model was optimised for deployment on the NVIDIA Jetson Nano GPU, capitalising on its high computational capacity and energy efficiency as shown in figure 5.5. To assure seamless operation on the Jetson Nano platform, the deployment procedure includes model optimisation techniques such as model pruning and quantization. Finally, the system is validated in real-world agricultural contexts, where the perfor-

mance and robustness of the deployed model are rigorously evaluated, confirming its practical utility and contribution to precision agriculture.

Table 5.5: Evaluation metrics of the fusion models with state-of-the-art pre-trained models

Model	Pixel Accuracy	IoU	Precision	Recall	F1-score	Parameters
Proposed Model	**0.91**	**0.84**	**0.90**	**0.92**	**0.91**	**3,39,394**
ResNet50	0.76	0.65	0.72	0.89	0.79	24,637,826
Inception	0.74	0.58	0.77	0.77	0.71	23,903,010
LeNet	0.82	0.70	0.79	0.94	0.82	61,326

5.2.4 Results Analysis

The U-Net model and proposed CNN of the fusion model are trained for 50 epochs. The collected dataset is partitioned as per the 10- Fold Cross Validation (10-FCV) for training and testing. Then training dataset is first feed into U-Net and trained with Adam optimizer with cross-entropy loss function. Next, output of U-Net is feed in to CNN model and trained using Adam optimizer with cross-entropy loss function. The training performance of U-Net and CNN is depicted in in figure 5.6 and figure 5.7 respectively. The individual binary masks of given petri dish generated by U-Net model are as shown in figure 5.8. These individual binary masks are combined to create binary mask of input image. These segmented images are used to extract the seeds as shown in figure 5.9. All the extracted seeds are shown in figure 5.11. These seeds are fed into the proposed CNN for classification of germination or no germination state.

5.2.5 10-Fold Cross Validation (10-FCV)

The 1O-FCV is employed to assess the efficacy of a deep learning model in predicting seed germination. Initially, the data set consisting of 1200 is partitioned into ten subsets of equal size 120. During each iteration, the training process utilizes nine subsets that is 1080 images for training purposes, while one subset of 120 images is exclusively reserved for testing. The model undergoes training using the training subsets and is subsequently assessed using the

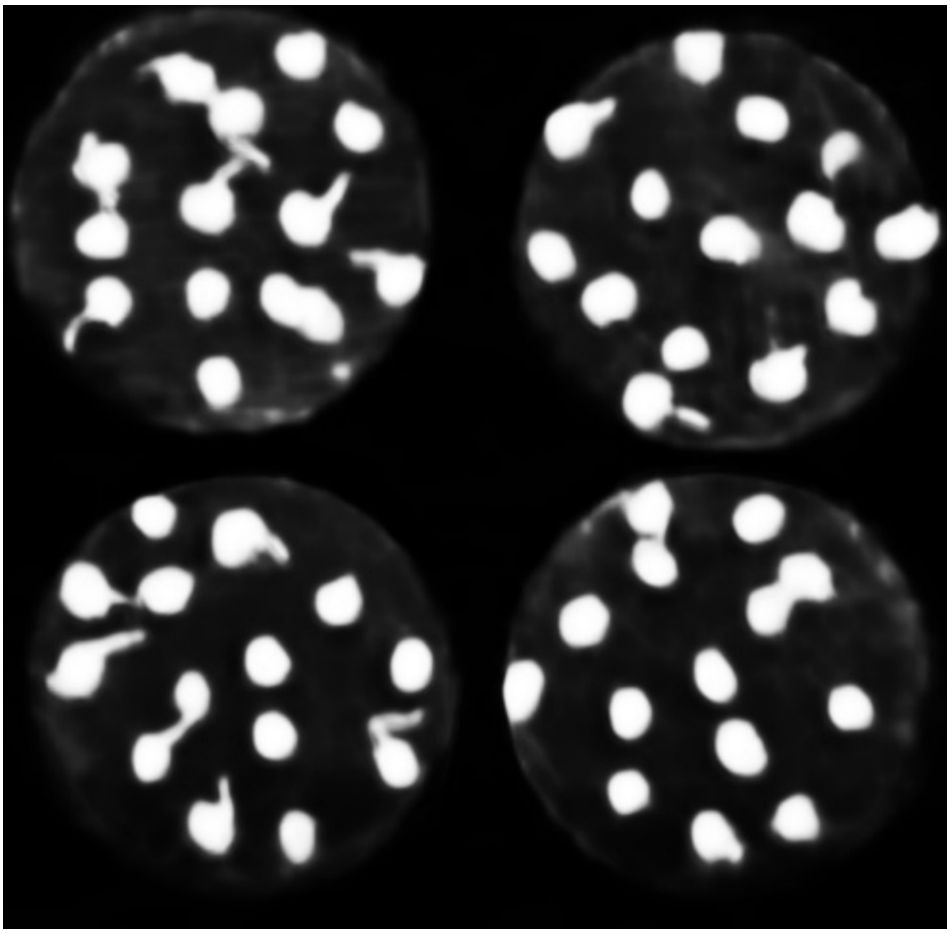


Figure 5.8: Binary masks of the seeds generated by U-Net model.

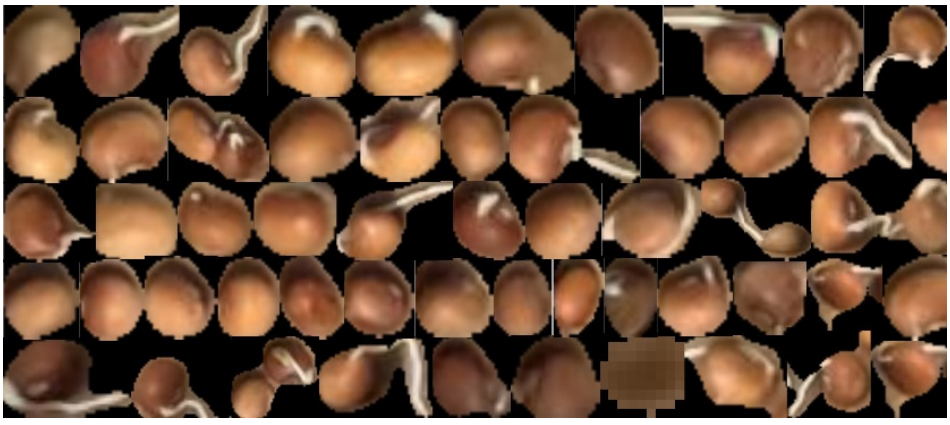


Figure 5.9: Seeds extracted from petri dishes using mask images of U-Net model.

validation subset. A computation determines an evaluation metric, such as error rate or mean absolute error. The procedure is executed ten times, with the validation subset of 120 images

being rotated after each iteration. Each individual data point is included in the validation set only once, guaranteeing the comprehensive utilization of all available data i.e. 1200 images. This particular methodology serves as a preventive measure against model over-fitting. The ten performance metrics are computed by taking their average, resulting in a single performance estimate as shown in table 5.4. This estimation offers a more comprehensive assessment of the model's efficacy.

5.2.6 Comparative Analysis

The proposed model is compared with three state-of-art models ResNet50, Inception, and LeNet respectively. These results are furnished in table 5.5 and best values are highlighted. From the table results it is observed that the proposed model performs better than the other three state-of-art models (ResNet50, Inception, and LeNet) overall. The suggested model predicts pixels with more accuracy (0.91 pixels accurate) than ResNet50 0.76, Inception 0.74, and LeNet (0.82 pixels accurate). The proposed fusion model has a score of 0.84 for Intersection over Union (IoU), which is much higher than ResNet50's (0.65), Inception's (0.58), and LeNet's (0.70) scores. This shows more overlap between the segmentation regions that were predicted and those that were actually segmented, further demonstrating the model's improved performance. The suggested model outperforms ResNet50 (0.72), Inception (0.77), and LeNet (0.79), achieving an accuracy score of 0.90 when measuring the model's ability to correctly identify relevant occurrences.Recall, a metric that gauges a model's capacity to find all relevant occurrences, gives the proposed model a score of 0.92. ResNet50 and Inception fall short with 0.89 and 0.77, respectively, although LeNet is near at 0.94.

Finally, the proposed fusion model outperforms ResNet50 0.79, Inception 0.71, and LeNet 0.82 with an F1 score of 0.91. This higher F1 score suggests that the suggested model offers a superior precision/recall balance, which is crucial in situations of class imbalance. In conclusion, the proposed fusion model outperforms the competition in terms of effectiveness and accuracy, making it the best option given these performance measures. Further, It is noted that the proposed model requires less number of parameters after LeNet. The model is carried

out the experiments with only U-Net and proposed a model for the germination task. The outcomes of these experiments are shown in the figure 5.12. The Proposed Fusion Model, which achieves an accuracy of 91.00% and consists of 19,62,625 trainable parameters, exhibits superior performance compared to both the UNET with Depthwise Separable Convolutions and the UNET with Atrous Spatial Pyramid Pooling. Consequently, it is deemed the most suitable choice for seed germination detection as shown in table 5.6.

Table 5.6: Evaluation Metrics of the fusion models with state of art models

Model	Accuracy	Parameters
UNET with Depthwise Seperable Convolutions	87.2	2,34,701
UNET with Atrous Spatial Pyramid Pooling	90.54	8,06,67,202
Proposed fusion model	**91.00**	19,62,625

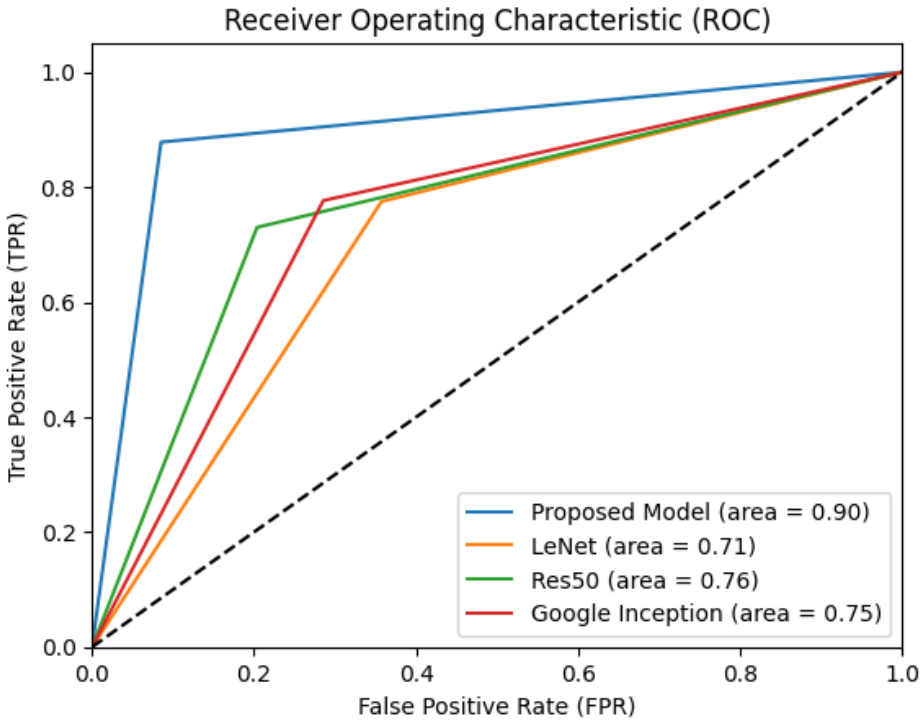


Figure 5.10: ROC curve for the proposed fusion model.

The ROC curve of the proposed model with state-of-the-art model is shown in figure 5.10. From the figure it is noted that the proposed model performance is consistently superior to those of the LeNet, Inception, and ResNet50 models. This would indicate that the proposed

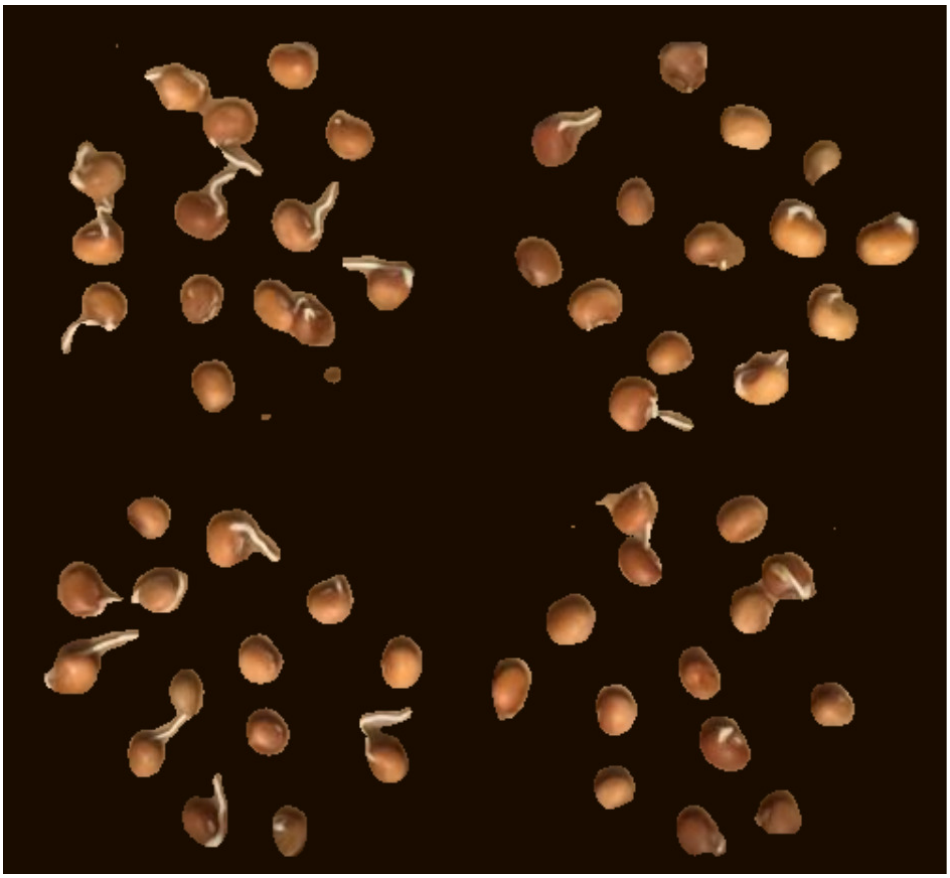


Figure 5.11: Collection of all the segmented masks from proposed fusion model.

model performs better with various threshold settings. The proposed model appears more capable of differentiating between positive and negative classes, resulting in a more reliable model for seed germination prediction.

5.2.7 Model Deployment on Jetson Nano

A U-Net model with a weight of 1.3 MB, and a CNN weight of 22KB, predict seed germination on the Jetson Nano in real time. A red bounding box indicates that the seed has not germinated, and a green bounding box indicates that it has germinated. We have trained our model using red gram seeds and tested with chickpea seeds and our model detected germination of chickpea seeds effectively as shown in figure 5.13 and 5.14. The U-Net's complicated design, which incorporates convolutional and max pooling layers in its contracting route and

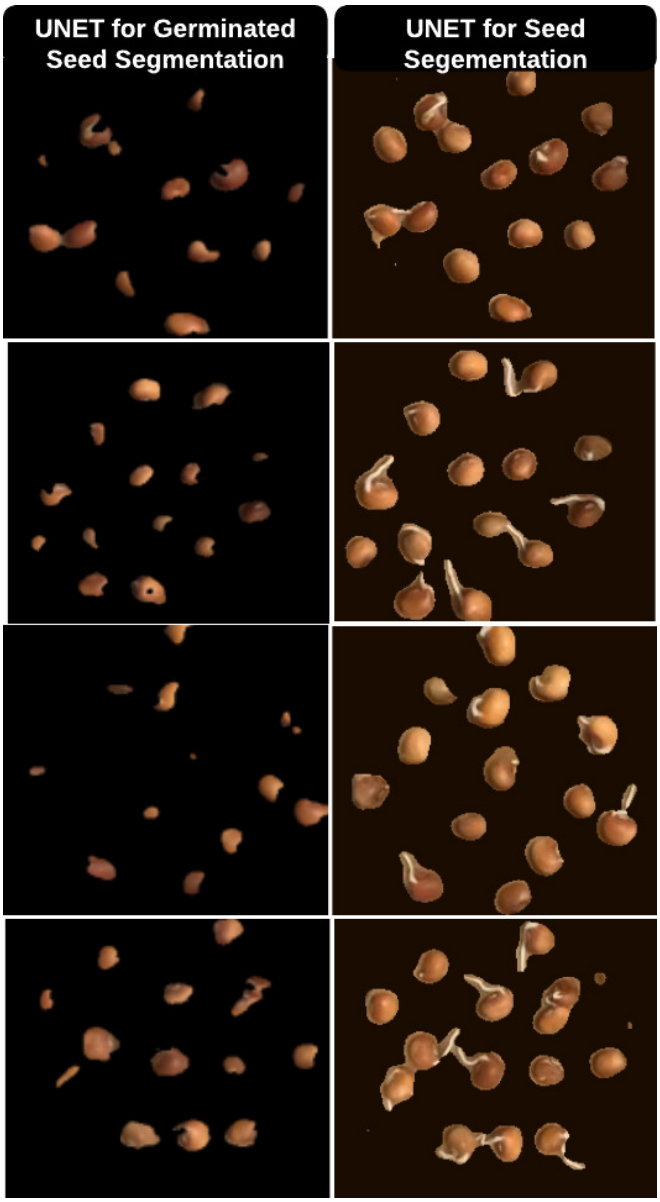


Figure 5.12: Comparision of U-Net only for germination and fusion model

an expanding path for exact localization, makes it more significant. The CNN model is lighter and quicker, but less precise.

These models continually analyze seed pictures to predict germination in less than 0.26 milli seconds. Real-time analysis might give farmers and researchers rapid feedback, enhancing their operations.

Though vital for its size, the Jetson Nano is resource limited. Thus, the implementation's

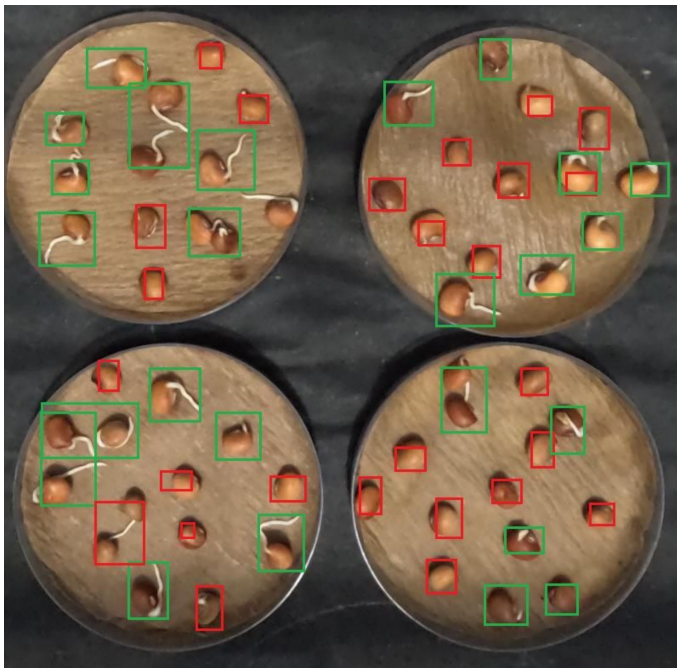


Figure 5.13: Germination prediction on Jetson nano hardware board.

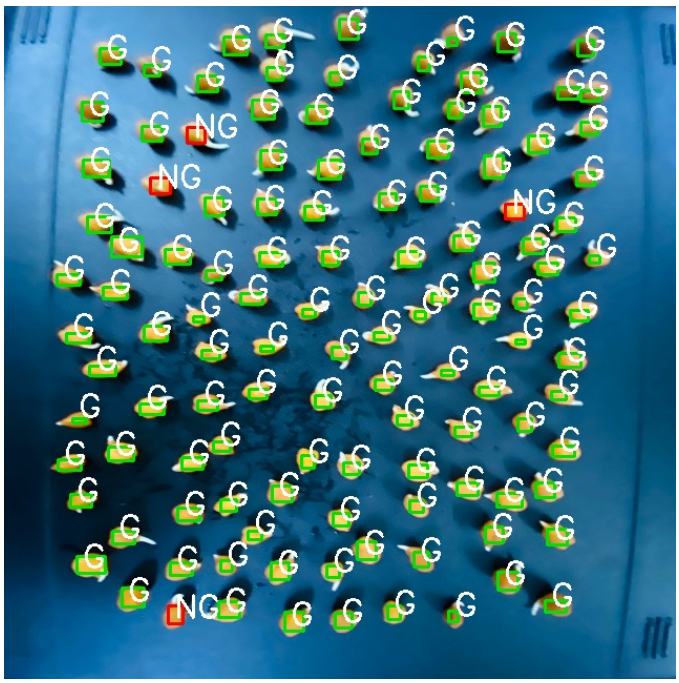


Figure 5.14: Germination prediction on Jetson nano board for 100 to 150 chickpea seeds.

efficiency depends on model complexity, picture resolution, and other parameters. Optimize your models and image pre-processing for device performance.

5.3 Summary

This chapter introduces a novel fusion model for seed detection and germination classification is proposed that combines the U-Net and CNN architectures. By harnessing U-Net's capabilities in image segmentation and CNN's strengths in classification, the proposed approach enables effective seed germination classification. Additionally, the model is specifically optimized for real-time processing by implementing it on the Nvidia Jetson Nano embedded GPU platform.The model is tested on real time environment and the measured latency is 0.26 milli seconds. The proposed fusion model obtained 0.91 pixel accuracy, 0.84 IoU, and 0.90 precision which are best when compared to state-of-the-art models. Also, the proposed model requires less number of trainable parameters after LeNet. Further, the proposed model is tested in real-time and achieved 0.26 ms latency.

Chapter 6

SeED-Net: Seed Encoding Decoding Network for Enhancing Seed Quality Analysis through Automation

In the previous chapter 5, we explored the application of the UNet architecture for the task of seed detection, which, despite its widespread use in image segmentation tasks, resulted in suboptimal accuracy levels for our specific application. This led us to the development of a customized Encoder-Decoder CNN architecture tailored specifically for the task at hand. Our novel approach builds on the foundational principles of CNNs but introduces several key modifications designed to enhance feature extraction and pattern recognition capabilities specifically for seed detection. Through a series of experiments, we systematically evaluated the performance of our custom-designed model against the UNet baseline, demonstrating significant improvements in detection accuracy. This chapter delves into the architectural nuances of our Encoder-Decoder CNN, including the rationale behind each design choice, the implementation details, and a comprehensive comparison of its performance against the UNet model. Our findings suggest that the customized Encoder-Decoder CNN not only surpasses the UNet in terms of accuracy but also offers insights into the potential for specialized architectures in addressing domain-specific challenges in image segmentation tasks.

6.1 Proposed Encoder Decoder

The Encoder Decoder CNNaaq model is specifically engineered for the task of image segmentation, with a particular focus on seed detection. This model is characterized by a thoughtful sequence of layers, including convolutional, pooling, and upsampling layers, complemented by strategic concatenation points. These elements collectively facilitate the model's ability to effectively process images and extract relevant features at various levels of abstraction, crucial for the accurate identification and segmentation of seeds within an image.

Table 6.1: Model summary of the proposed SeED-Net

Layer Type	**Num. of Filters**	**Kernel Size**	**Activation**
InputLayer	-	-	-
Conv2D	64	3x3	ReLU
MaxPooling2D	-	2x2	-
Conv2D	128	3x3	ReLU
MaxPooling2D	-	2x2	-
Conv2D	256	3x3	ReLU
MaxPooling2D	-	2x2	-
Conv2D	512	3x3	ReLU
MaxPooling2D	-	2x2	-
Conv2D	1024	3x3	ReLU
Conv2DTranspose	512	2x2	ReLU
Concatenate	-	-	-
Conv2D	512	3x3	ReLU
Conv2DTranspose	256	2x2	ReLU
Concatenate	-	-	-
Conv2D	256	3x3	ReLU
Conv2DTranspose	128	2x2	ReLU
Concatenate	-	-	-
Conv2D	128	3x3	ReLU
Conv2DTranspose	64	2x2	ReLU
Concatenate	-	-	-
Conv2D	64	3x3	ReLU
Conv2D	2	1x1	Softmax

6.1.1 Proposed Encoder Decoder Structure

The architecture commences with an input layer designed to accept images of a predefined shape, marking the start of the encoder phase. The proposed SeED-Net model summery is shown in table 6.1. Here, the model employs convolutional layers to apply filters for feature extraction and pooling layers to reduce the spatial dimensions of the resulting feature maps, all while preserving essential information. This phase is crucial for distilling the image into a form that highlights the most relevant features for seed detection. As the model progresses, an increase in the number of filters within convolutional layers ensures the capture of a broad spectrum of features, ranging from basic textures to complex patterns indicative of seed forms.

Transitioning through the "Center Block," which serves as the core of the feature encoding process, the model then shifts into the decoder phase. This segment is pivotal for segmentation tasks, aiming to accurately outline the contours of seeds. Upsampling layers are utilized to incrementally enlarge the feature maps back to the input image size, while concatenation layers merge these upsampled features with those from the encoding phase. This approach ensures that the model benefits from both high-level semantic insights and detailed textural information, enabling precise segmentation.

6.1.2 Advantages Over U-Net

The adaptation and fine-tuning of this model for seed detection have resulted in performance that surpasses the well-regarded U-Net architecture in similar tasks. Several aspects contribute to this enhanced effectiveness:

6.1.2.1 Layer Parameter Optimization

The model's configuration, including the use of ReLU activation functions, addresses common training challenges like vanishing gradients, facilitating efficient learning.

6.1.2.2 Balanced Feature Processing

The symmetric design, incorporating both downsampling and upsampling with feature concatenation, ensures the preservation of crucial information throughout the model's layers.

6.1.2.3 Customized Network Depth and Filters

Adjustments to the depth of the network and the specifications of convolutional filters allow for a nuanced analysis of seed images, capturing detailed features essential for differentiating seeds in complex visuals.

6.1.3 Tailoring for Precision

The architecture's success in seed detection illustrates the importance of customizing neural network designs to meet the unique demands of specific applications. By balancing deep feature extraction with detailed reconstruction, the model achieves segmentation with high precision, outperforming established standards like the U-Net. This advancement not only highlights the model's capability in seed detection but also opens avenues for further customizations in image segmentation tasks across various domains.

6.1.4 Mathematical Foundations of Encoder and Decoder CNN

6.1.4.1 Convolutional Layers

The convolution operation for layer i is shown in equation 6.1

$$F_i(x, y) = \sum_{m=-a}^{a} \sum_{n=-b}^{b} I(x - m, y - n) \cdot K(m, n) \tag{6.1}$$

where $F_i(x, y)$ represents the feature map, $I(x, y)$ is the input image or feature map from the previous layer, and $K(m, n)$ is the convolutional kernel.

6.1.4.2 Pooling Layers

Following convolution, max pooling is applied to reduce dimensionality as shown in equation 6.2

$$P_i(x, y) = \max_{(m,n)\in W} F_i(x + m, y + n) \tag{6.2}$$

where $P_i(x, y)$ is the pooling layer output, and W denotes the pooling window.

6.1.4.3 Upsampling Layers

The upsampling process is aimed at restoring spatial dimensions is shown in equation 6.3

$$U_i(x, y) = F_i(\lfloor x/s \rfloor, \lfloor y/s \rfloor) \tag{6.3}$$

where s is the scale factor for upsampling.

6.1.4.4 Activation Functions

The ReLU activation function introduces non-linearity as shown in equation 6.4

$$R(x) = \max(0, x) \tag{6.4}$$

6.1.4.5 Concatenation in Decoder

Concatenation merges encoder features with upsampled features as showin in equation 6.5

$$C_i = \text{concat}(F_{\text{encoder}}, U_i) \tag{6.5}$$

6.1.5 Transposed Convolution

Transposed convolution, also known as fractionally-strided convolution, is a key operation often used in semantic segmentation tasks in deep learning. It helps in reversing the down-sampling operations that occur during regular convolutions, making it possible to restore the spatial dimensions of the feature maps. The goal of transposed convolution is to produce an

output with spatial dimensions that match those of the input, which is particularly important in tasks like pixel-level classification in semantic segmentation.

6.1.6 Basic Operation

Ignoring channels for now, the basic transposed convolution operation involves sliding a kernel window with a stride of 1 over the input matrix. In this process, imagine each element in the input matrix shaking hands with its counterpart in the kernel. They then engage in a little dance, where each pair multiplies their values together. After the dance, all the results are gathered and added up to form the output matrix. It's like each element in the input matrix teams up with its partner in the kernel to contribute to the final result, creating a synchronized performance that yields the output matrix.

The intermediate results are computed by multiplying elements of the input matrix by the kernel and are summed to produce the final output.

The mathematical calculation involves the element-wise multiplication of the input matrix and the kernel, followed by summation to obtain the output matrix. When the input and kernel are both four-dimensional tensors, high-level APIs can be used to perform the transposed convolution.

6.1.7 Padding, Strides, and Multiple Channels

In transposed convolution, padding is applied to the output rather than the input, with the specified padding number affecting the dimensions of the output. Strides are applied for intermediate results and the output rather than the input, and changing the stride affects the size of the intermediate and output tensors.

Additionally, for multiple input and output channels, the transposed convolution operation works similarly to regular convolution. When specifying multiple output channels, there is one kernel for each output channel, similar to the regular convolution.

6.1.8 Connection to Matrix Transposition

The name "transposed convolution" is connected to matrix transposition. Convolutions can be implemented using matrix multiplications, where the weight matrix is obtained from the convolutional kernel. In the transposed convolution layer, both forward and backpropagation functions can be carried out by multiplying the input vector with the weight matrix and its transpose, respectively. This means that during forward propagation, the input vector interacts with the weight matrix to produce the output, while during backpropagation, the input vector interacts with the transpose of the weight matrix to calculate the gradients necessary for adjusting the weights.

6.2 Weight Initialization

Weight initialization is a critical consideration in the design and training of deep learning neural network models. This process entails establishing the initial values for the parameters, such as weights, in the neural network before commencing the training of models on a dataset. The selection of appropriate weight initialization techniques can significantly impact the effectiveness and efficiency of the optimization process during training.

The historical approach to weight initialization typically involved using small random values within certain ranges, such as $[-0.3, 0.3]$ or $[-1, 1]$. However, over the last decade, more specific heuristics have been developed based on extensive research into the behavior of activation functions and the number of inputs to the nodes. These specialized heuristics have become the de facto standard due to their capability to optimize the training process and improve the performance of neural network models.

Modern weight initialization techniques are categorized based on the type of activation function used in the network nodes, such as Sigmoid or Tanh, and Rectified Linear Unit (ReLU). For nodes employing Sigmoid or Tanh activation functions, the Xavier and normalized Xavier weight initialization heuristics are commonly used. On the other hand, the He weight initialization heuristic is tailored for nodes using the ReLU activation function, fre-

quently employed in hidden layers of multilayer perceptron and convolutional neural network models.

6.2.1 Xavier Weight Initialization

Xavier initialization in figure 6.1 involves picking random numbers from a certain range, specifically between $-(1/\sqrt{n})$ and $1/\sqrt{n}$, where n stands for the number of inputs that go into a particular node. This technique ensures that the weights are appropriately scaled to account for the number of inputs, thereby promoting more stable and effective learning. It is particularly suitable for nodes using Sigmoid or Tanh activation functions.

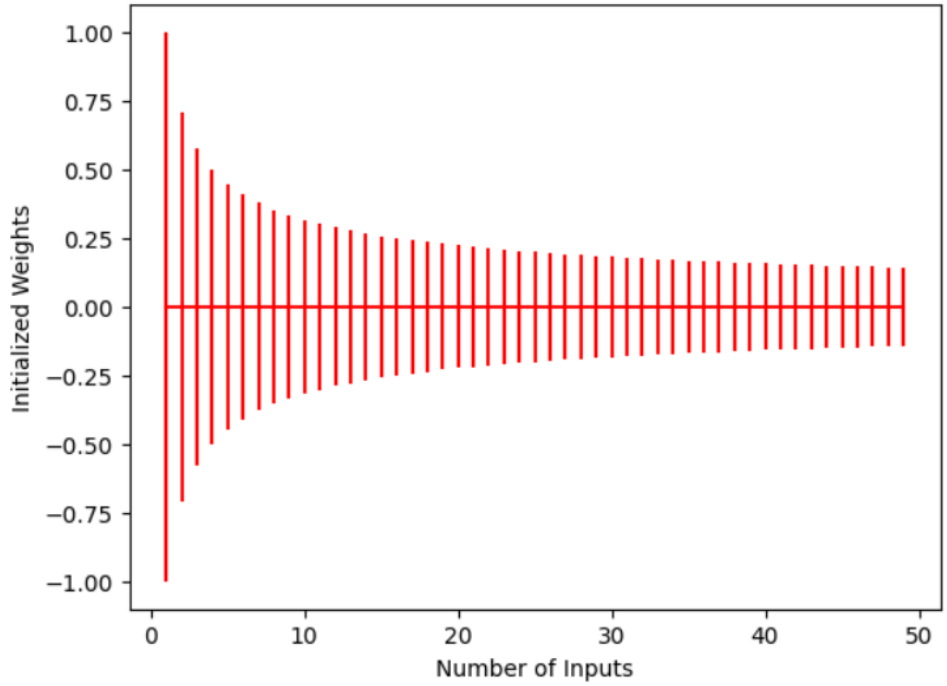


Figure 6.1: Xavier Weight Initialization

6.2.2 Normalized Xavier Weight Initialization

The normalized Xavier initialization method shown in figure 6.2 extends the Xavier approach by incorporating the number of outputs from the layer in addition to the number of inputs. It calculates the weights using a uniform probability distribution within the range

$\{-(\sqrt{6}/\sqrt{n+m}), \sqrt{6}/\sqrt{n+m}\}$, where n signifies the number of inputs to the node and m denotes the number of outputs from the layer. This technique aims to maintain a balanced range of weights, further enhancing the stability and convergence of the model during training.

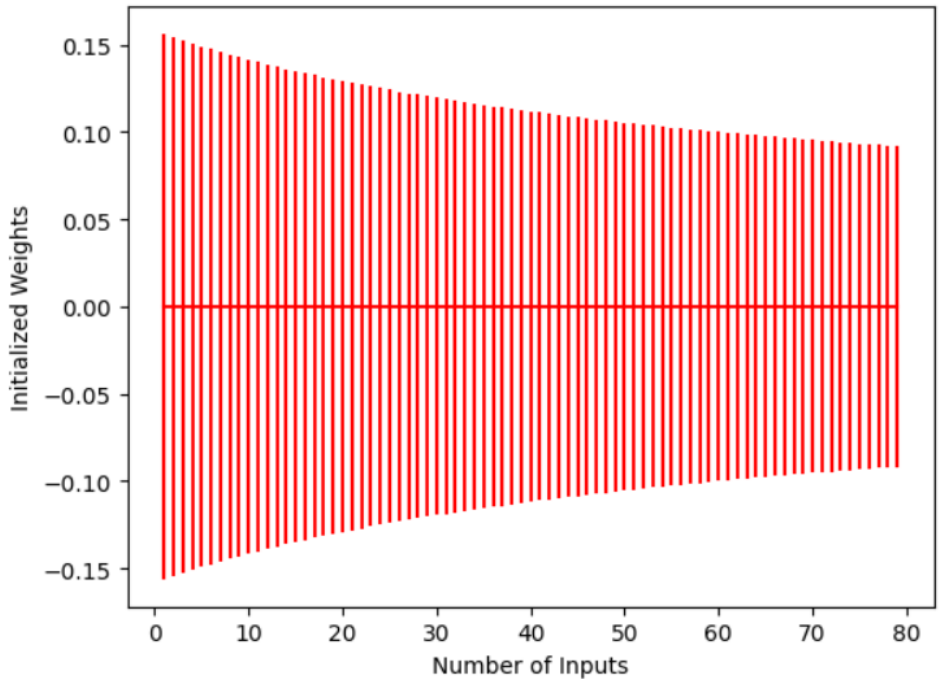


Figure 6.2: Normalized Xavier weight initialization

6.2.3 He-Weight Initialization

The He initialization method shown in figure 6.3, named after Kaiming He, is designed for layers employing the ReLU activation function. It involves initializing weights using a Gaussian probability distribution with a mean of 0.0 and a standard deviation of $\sqrt{2/n}$, where n represents the number of inputs to the node. This technique addresses issues associated with the dying ReLU problem and ensures that the initial weights align with the characteristics of ReLU activations, leading to more effective and stable training.

Overall, weight initialization is a crucial step in the development of neural network models, and the adoption of modern heuristics, such as Xavier, normalized Xavier, and He weight

initialization, has significantly contributed to the improved training and performance of deep learning models.

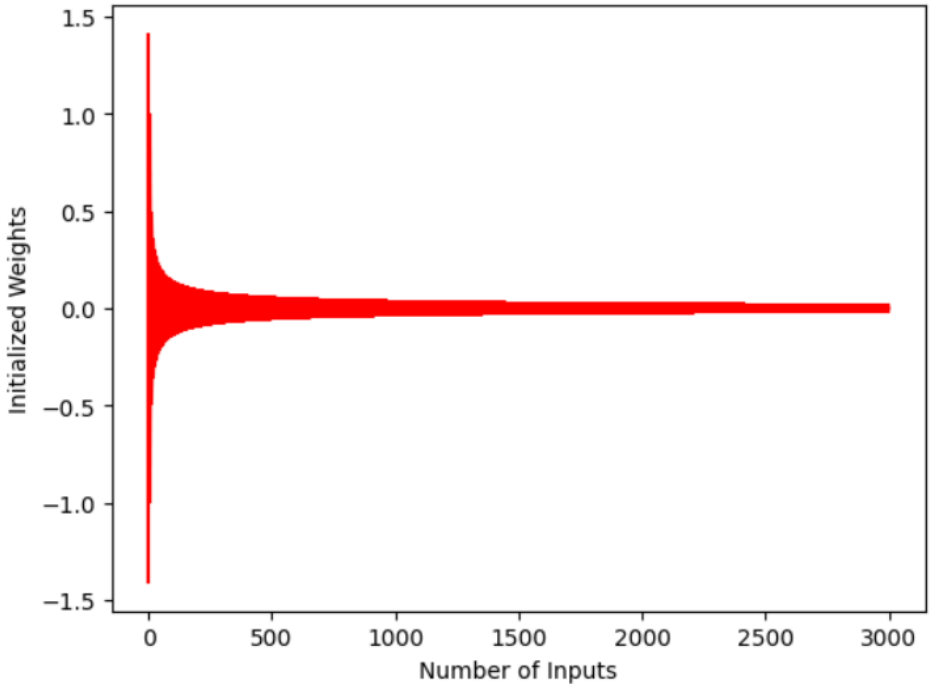


Figure 6.3: He weight Initialization

6.3 Enhanced Performance over U-Net

The model demonstrates enhanced seed detection capabilities, outperforming established methods such as U-Net.

The proposed Encoder Decoder CNN model's success in seed detection underscores the potential of tailored architectures and precise mathematical modeling in addressing complex image segmentation tasks. This approach opens new avenues for research, suggesting the model's adaptability to a wide range of segmentation challenges.

6.4 Proposed Methodology

The Seed Encoding-Decoding Instant Segmentation Network (SeED-Net) is a deep learning architecture tailored for accurate seed segmentation and germination classification. It utilizes an encoder-decoder approach to perform pixel-level analysis on seed datasets using an embedded GPU. Figure 6.4 shows the architecture of SeED-Nework is a sophisticated deep learning framework designed for the precise tasks of segmenting seed images and classifying their germination status. The architecture, as depicted in the referenced figure, is built on a combination of encoding, decoding, and classification layers tailored specifically for the analysis of seed images.

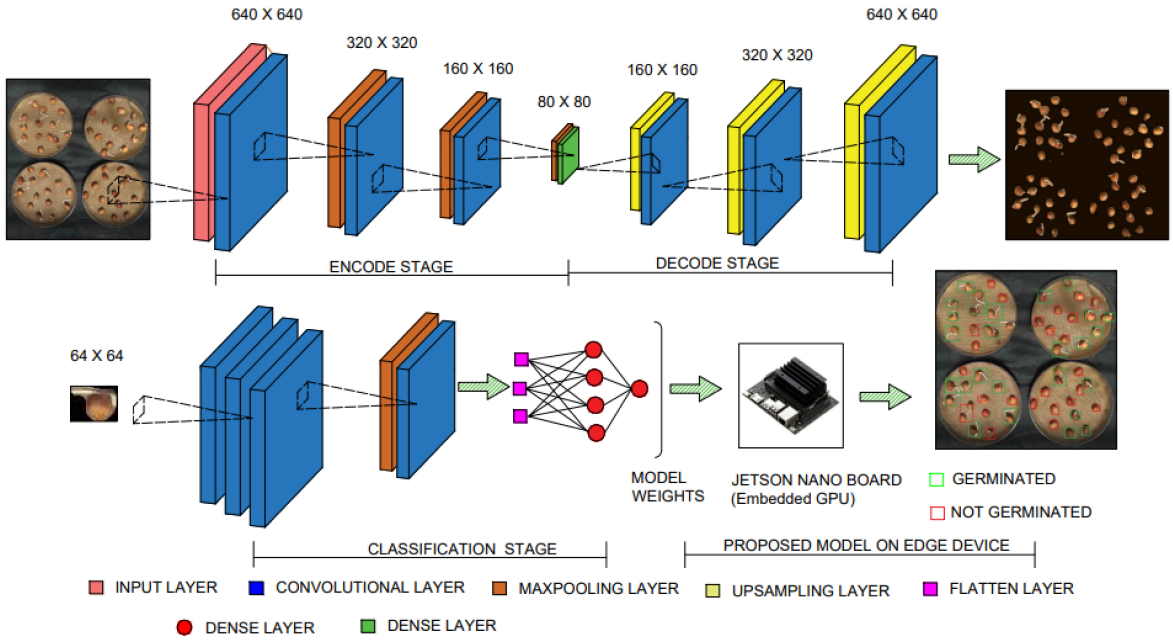


Figure 6.4: Proposed seed encoding decoding instant segmentation and germination classification network

6.4.1 Instant Segmentation Network

As shown in table 6.2 the input layer is the first layer of a neural network. It receives the initial input data and passes it on to the next layer. The input shape of the data is represented as (None, 640, 640, 3). The input layer receives images of seeds of dimensions (640, 640, 3),

where the value 3 denotes the three RGB channels. The convolutional layer is a fundamental component of CNNs. It plays a crucial role in extracting meaningful features. The output shape of the model is (None, 640, 640, 32). The number of trainable parameters are 896. The present layer is a 2D convolutional layer that consists of 32 filters and utilizes a kernel size of (3, 3) to perform convolution on the input picture. Next layer is the Max Pooling Layer is a commonly used component in CNNs for feature extraction. It operates by partitioning the input feature map into not . The shape of the data is represented as (None, 320, 320, 32). The next layer is a max-pooling layer that effectively decreases the spatial dimensions of the preceding feature maps by a factor of two through the use of 2x2 pooling. The output shape of the layer is (None, 320, 320, 64). The trainable parameters are 18,496.

The subsequent layer consists of a 2D convolution operation with 64 filters and a kernel size of (3, 3) applied to the existing feature maps. The shape of the input data is (None, 160, 160, 64). The subsequent layer is a max-pooling layer that further decreases the spatial dimensions. The output shape of the model is (None, 160, 160, 128). The parameters provided for analysis are 73,856. The feature maps are subjected to a 2D convolutional layer with 128 filters and a kernel size of (3, 3). The subsequent layer is a max-pooling layer, which serves to decrease the spatial dimensions. The max pooling layer is used to downsample the input feature maps, reducing their spatial dimensions while retaining them. Output the shape of the data is (None, 80, 80, 128). The dense layer, also known as the fully connected layer, is a type of layer commonly used in neural networks. The term "dense" refers to a state or quality of being closely compacted or crowded together. It is often used to change the output dimensionality and classify seed images based on output from convolutional layers. The shape of the data is (None, 80, 80, 128). The trainable parameters are 16,512.

A linked layer with a total of 128 units is used. The UpSampling2D layer is a component used in deep learning models to increase an input tensor's spatial dimensions. The output shape of the model is (None, 160, 160, 128). The upsampling layer is designed to enhance the spatial dimensions by a multiplication factor 2. The CNN output shape of the model is (None, 160, 160, 128). The parameter value provided is 147,584. The upsampled feature maps are subjected to a 2D convolutional layer with 128 filters and a kernel size of (3, 3).

Table 6.2: Seed encoding-decoding instant segmentation network Architecture

Layer (Type)	**Output Shape**	**Param**
Input layer	(640,640,3)	
Convolutional layer	(640,640,32)	896
Max Pooling layer	(320,320,32)	
Convolutional layer	(320,320,64)	18,496
Max Pooling layer	(160,160,64)	
Convolutional layer	(160,160,128)	73,856
Max Pooling layer	(80,80,128)	
Dense layer	(80,80,128)	16,512
Upsampling layer	(160,160,128)	
Convolutional layer	(160,160,128)	1,47,584
Upsampling layer	(320,320,128)	
Convolutional layer	(320,320,64)	73,792
Upsampling layer	(640,640,64)	
Convolutional layer	(640,640,1)	577
	Total	**3,31,713**

CNN output shape is (None, 320, 320, 128). The subsequent layer is an upsampling layer that further augments the spatial dimensions and it plays a crucial role in extracting and learning. The output shape of the model is (None, 320, 320, 64).

The parameters provided for analysis are 73,792. The upsampled feature maps are subjected to a 2D convolutional layer with 64 filters and a kernel size of (3, 3). The UpSampling2D layer is a component used in deep learning models for upsampling or increasing the spatial dimensions of an input tensor. Output The shape of the data is represented as (None, 640, 640, 64). The subsequent layer is an upsampling layer that further enhances the spatial dimensions. The convolutional layer is a fundamental component of CNNs used for image processing and pattern recognition tasks. The output shape of the model is (None, 640, 640, 1). Trainalbe parameters are 577. The seed image reconstruction process involves the utilization of a final 2D convolutional layer. This particular layer consists of a single filter, resulting in the generation of a single channel. The kernel size employed in this convolutional layer is (3, 3). The total trainable parameters are 3,31,713 for the proposed model.

6.4.2 Seed Germination Classification Using Proposed CNN Model

CNN is provided with high-resolution pictures of seedlings at different stages of germination. Each image may potentially represent the initial stages of seed sprouting, a seed that has fully germinated, or a seed that exhibits no visible indications of germination. The convolutional layer is a fundamental component of CNN. It is responsible for extracting features from input data by applying a series of convolutional filters. In this layer, filters are applied to the seed image in order to identify fundamental characteristics, including seed outlines, textures, and the early appearance of a radicle or shoot. From a mathematical perspective, the process involves capturing distinct germination patterns by the multiplication of matrix values with the corresponding picture segment, followed by the summation of these outcomes.

The pooling layer is a crucial component in CNN. It is responsible for reducing the spatial dimensions of the input feature maps, hence decreasing the computational complexity. By decreasing the dimensions of the feature map, the model is able to focus on important elements such as the development direction of a sprouting seed or the differentiation between seed and soil, hence improving computing efficiency. Fully connected layers, also known as dense layers, are a fundamental component of artificial neural networks. These layers consist of nodes, or neurons, that are connected to every node in the previous layer. During this phase, a thorough analysis is conducted on the prominent characteristics of the seed, such as the sprouting root's prominence and the swelling of the seed. The aforementioned characteristics are compressed into a structure that is appropriate for the ultimate categorization, enabling the anticipation of whether the seed is undergoing germination or not.

The flattened layer inside a CNN functions as an intermediary component connecting the convolutional layers and the fully connected layers. In the case of seed germination classification, the process involves categorizing seeds into two classes i.e. germinated or not germinated. Following the extraction of features from convolutional layers, the flatten layer is utilized to transform the multi-dimensional feature maps into a one-dimensional vector. The flattened vector effectively preserves crucial information about the seed's condition and facilitates data processing by the dense layers. In this context, the network assumes the ul-

timate responsibility of deciding the classification of the seed by analyzing the discovered traits, therefore establishing whether the seed is germinated or remains in a dormant state. As shown in the table 6.3, total trainable parameters are 52,162 for the proposed classification model. Semantic segmentation's granular and pixel-wise methodology in categorizing seed germination offers a comprehensive analysis of even the slightest variations in the seed picture. The acquisition of high-resolution information is of utmost importance because the initial phases of germination may present as inconspicuous visual indications, such as minuscule protrusions or subtle alterations in coloration within seed areas. Without segmentation, it may not effectively capture or potentially homogenize these slight, geographically specific alterations. Consequently, in a non-segmented study, there is a risk of overlooking or disregarding early germination indications, which are of utmost importance for prompt interventions or evaluations. By neglecting the comprehensive perspective provided by segmentation, there is a potential drawback of failing to encompass the entirety of a seed's germination process.

Table 6.3: Proposed SeED-Net Classification Model

Layer (Type)	**Output Shape**	**Param**
Convolutional layer(Conv2D)	multiple	896
Maxpooling layer	multiple	0
Convolutional layer	multiple	18,496
Maxpooling layer(maxpooling 2d)	multiple	0
Flatten layer	multiple	0
Dense layer	multiple	32,770
	Total	**52,162**

6.5 Encoder-Decoder Network

6.5.1 Embedded Lookup

In the context of seed image analysis with Neural Machine Translation (NMT), input images are represented as high-dimensional arrays of pixel values. Embedded lookup is the process of transforming these pixel values into continuous vector representations, known as

embeddings. These embeddings capture essential visual features and patterns present in the seed images. Through the learning process, the embeddings are optimized to encode relevant information, enabling the model to effectively analyze seed images and extract meaningful insights.

6.5.2 Encoder-Decoder Architecture

The encoder-decoder architecture serves as the foundational framework for NMT models applied to seed image analysis. In this architecture, the encoder processes the input seed images and produces a condensed representation, often referred to as a context vector. This context vector encapsulates the essential features and characteristics of the input images. Subsequently, the decoder utilizes this context vector to generate output predictions, such as segmentation masks or classification labels, corresponding to different aspects of seed morphology or growth status. Through training, the encoder and decoder are jointly optimized to accurately capture the relationships between input seed images and their corresponding analyses.

6.5.3 Softmax Function

The softmax function plays a critical role in the decoder component of NMT models for seed image analysis, particularly in generating output predictions. After processing the context vector, the decoder predicts the probability distribution over different classes or labels relevant to seed analysis tasks. The softmax function is applied to these raw prediction scores, converting them into probabilities. These probabilities indicate the likelihood of each class or label being present in the input seed image. The class or label with the highest probability is then selected as the final prediction. By leveraging the softmax function, the decoder can produce reliable and accurate predictions, enabling robust analyses of seed images.

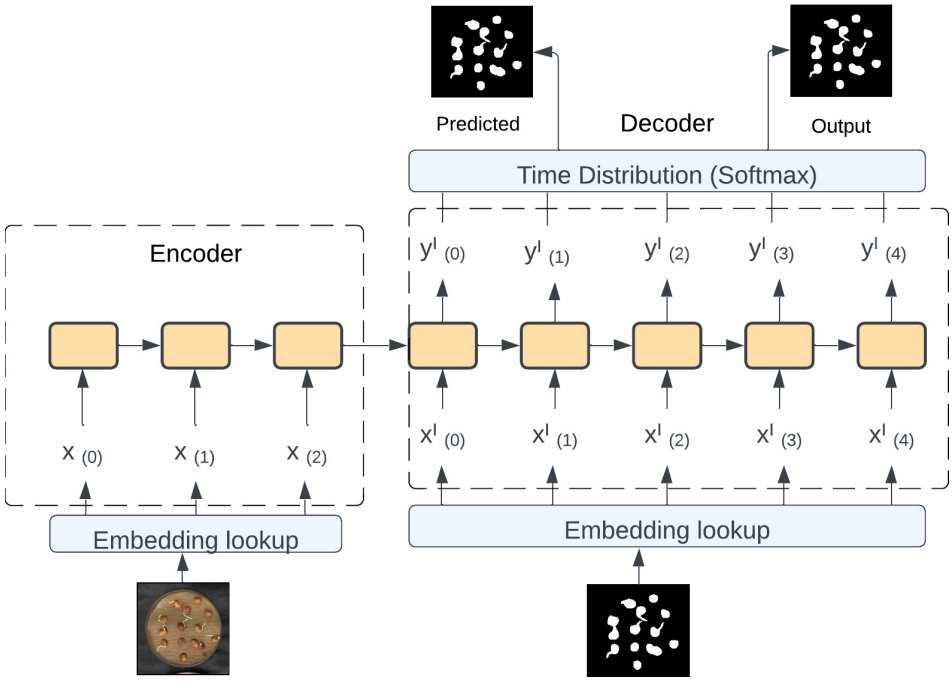


Figure 6.5: Process flow of encoder and decoder architecture for seed images

6.6 Experimental Results

6.6.1 Experimental Setup

The experimental setup utilized to acquire seed images includes a Jetson Nano with a camera interface. The process of capturing high-resolution seed photos is automated and occurs at regular intervals of five minutes. This systematic approach ensures the availability of a comprehensive and up-to-date dataset that accurately depicts seed presence and location dynamics. Moreover, the dataset accurately captures seed images inside a controlled environment by maintaining optimal humidity and temperature.

The Jetson Nano is connected to the MIPI CSI camera module for interface purposes. Python is utilised for the execution of image access and management tasks, with OpenCV being used for this purpose. A Python script was developed to automate the process of capturing images at regular intervals of five minutes. The proposed system regularly captures seed images in high definition, as demonstrated in figure 6.6. Figure 6.6 (a) of the figure shows a camera integrated with the Jetson Nano AI board, a key component for capturing images

for our dataset. Figure 6.6 (b) illustrates the operating system configured on a memory card, along with the external ports of the Jetson Nano, showcasing the complete setup required for our image processing tasks. To facilitate categorization and facilitate convenient retrieval, each image is accompanied by a timestamp and thereafter stored in an archival system, as illustrated in the figure 6.7. Prior to being inputted into the SeED-Ne for further analysis, the collection of seed images, total 35,871, undergoes a sequence of preprocessing processes. In the beginning, a large dataset consisting of 35,871 seed images are acquired, encompassing a wide variety of seed conditions and types. Prior to being fed into the model, the images undergo a sequence of preprocessing processes.Image resizing is a process in which all images are adjusted to adhere to a standardised dimension i.e 640x640. This practice ensures consistency and facilitates compliance with the input layer of the model. The preprocessed images are subsequently inputted into the proposed SeED-Ne, an exquisitely designed deep learning framework specifically tailored for the purpose of semantic segmentation tasks. The model analyses the seed images, extracting complex traits and patterns that are unique to the conditions of the seeds. The performance evaluation of the Encoder-Decoder model was conducted utilising the Keras and Tensor Flow 2.12.0 frameworks on an HP Z6 G4 Workstation, which is equipped with 52 cores and 64 GB of DDR4 RAM. The proposed Encoder-Decoder architecture utilized in this study was trained for seed detection using semantic segmentation. Proposed classification for seed germination prediction.The training process involved using a dataset consisting of 35,871 seed images and was conducted across 50 epochs. The training was performed on a GPU for efficient computation. Similarly, the CNN architecture under consideration was subjected to a training process spanning 50 epochs.

After undergoing rigorous training on the HP Z6 G4 GPU workstation, the SeED-Net model's weights are subsequently loaded to the Jetson Nano. This feature allows the compact and high-performance device to perform real-time prediction regarding seed detection and germination. The process of transferring weights to the Jetson Nano device enables real-time analysis at the location, offering significant value for applications that necessitate prompt feedback. This is particularly advantageous in agricultural contexts where quick assessments of seed viability hold utmost importance. By utilising the computing capabilities of the Jetson

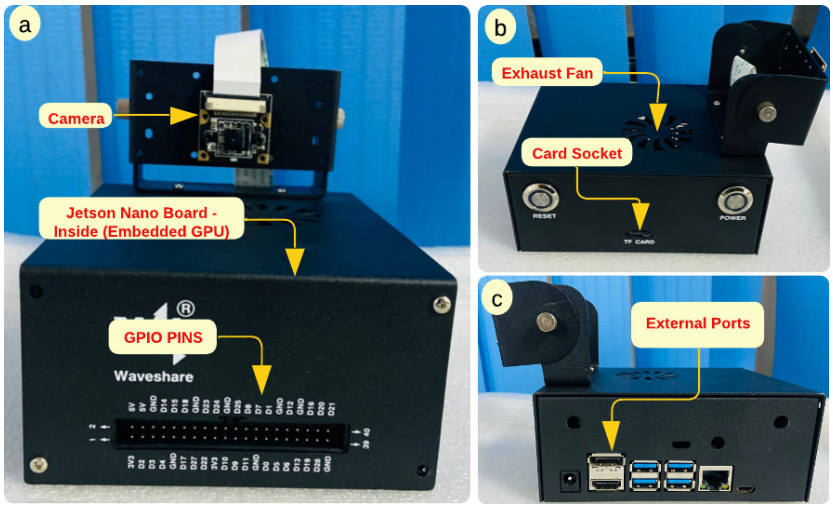


Figure 6.6: Jetson nano AI board (a) Camera intergated to Jetson Nano (b) Operating system ported on to memory card and external ports of jetson nano

Nano alongside the optimised weights of the SeED-Net, the forecast accuracy and speed of seed germination may be significantly improved.

6.6.2 Evaluation Metrics

When assessing the effectiveness of SeED-Net, a series of metrics used are as follows:

6.6.2.1 Pixel Accuracy

As shown in equation 6.6, pixel accuracy is a statistic that directly measures the proportion of correctly predicted pixels to the total number of pixels. Despite being computationally and conceptually straightforward, its limitation rests in its inability to offer comprehensive insights into individual class forecasts. The statistic in question can occasionally be misleading, particularly in situations where there are disparities in class distribution.

$$PixelAccuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{6.6}$$

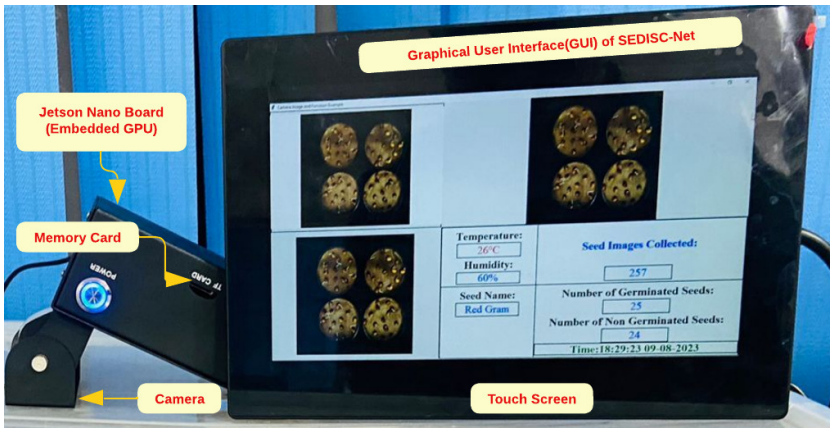


Figure 6.7: Proposed SeED-Net model using Jetson nano for real time germination prediction

6.6.2.2 Intersection over Union (IoU)

It is a metric commonly used in the field of computer vision to evaluate the accuracy of object detection algorithms. The IoU, as defined in equation 6.7, quantifies the degree of overlap between the anticipated segmentation of the model and the ground truth segmentation. The IoU metric quantifies the precision of a model by calculating the ratio of the intersection area to the union of the two areas. A higher IoU value shows that SeED-Net performs with greater accuracy in its segmentation jobs.

$$IoU = \frac{TP}{TP + FP} \tag{6.7}$$

6.6.2.3 Precision

The precision metric measures the proportion of accurate positive predictions out of all the expected positive instances. The proficiency of SeED-Net in accurately identifying pertinent segments is demonstrated by this indication.

6.6.2.4 Recall

As indicated by equation 6.8, quantifies the ratio of accurately detected positive instances to the total number of genuine positive segments. SeED-Net's capacity to regularly identify

noteworthy portions is a testament to its efficacy.

$$Recall = \frac{TP}{TP + FN} \tag{6.8}$$

6.6.2.5 F1 Score

The F1 score, as depicted in equation 6.9, combines precision and recall in a balanced manner, providing a comprehensive metric that is particularly useful in situations when there is an unequal distribution of classes. A higher F1 score indicates that the performance of SeED-Ne is improved in terms of accuracy.

$$F1Score = \frac{2TP}{2TP + FP + FN} \tag{6.9}$$

The utilisation of these measures guarantees a comprehensive assessment of SeED-Net, which is essential for the refinement and enhancement of its performance in the task of seed segmentation.

6.7 Results

The SeED-Net architecture, consisting of an encoder-decoder structure and a dedicated CNN component, is trained for 50 epochs. The dataset is partitioned according to the 10-FOLD cross-validation (10-FCV) technique, which guarantees the integrity of the training and testing stages. The training dataset is first subjected to processing through the encoder-decoder component of SeED-Net. This processing involves training the model using the Root Mean Square Propagation (RMSProp) with a cross-entropy loss function. Following this, the model's output is utilized as the input for the convolutional neural network (CNN) component of SeED-Net. The CNN is trained using theRMSProp optimizer with a cross-entropy loss function. The training evolution of the encoder-decoder segment and the CNN component is depicted in figure 6.8 correspondingly. The individual layers output generated during the encoder-decoder phase are depicted in figure 6.9. The unique masks combine to create a

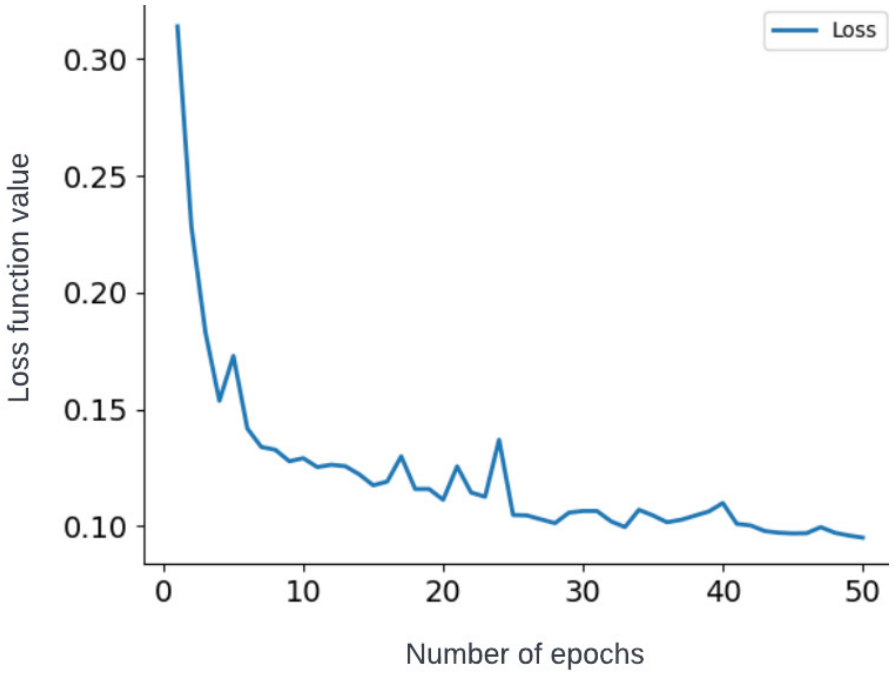


Figure 6.8: Loss curve for the proposed SeED-Net system

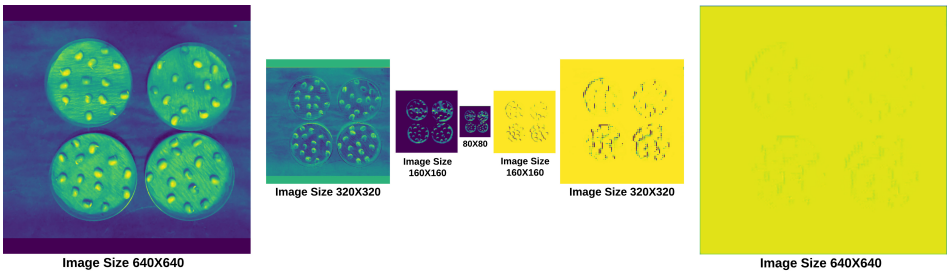


Figure 6.9: Layer outputs of proposed SeED-Nework

comprehensive binary mask of the original image, as shown in figure 6.10. The comprehensive compilation of individual seeds is depicted in figure 6.10. Binary images generated by the proposed model is shown in figure 6.11.Segmented images of the proposed system are shown in figure 6.12. The seeds that have been carefully extracted are subsequently sent to the CNN component of SeED-Net to determine their germination status, which may be classified as either germinated or non-germinated.

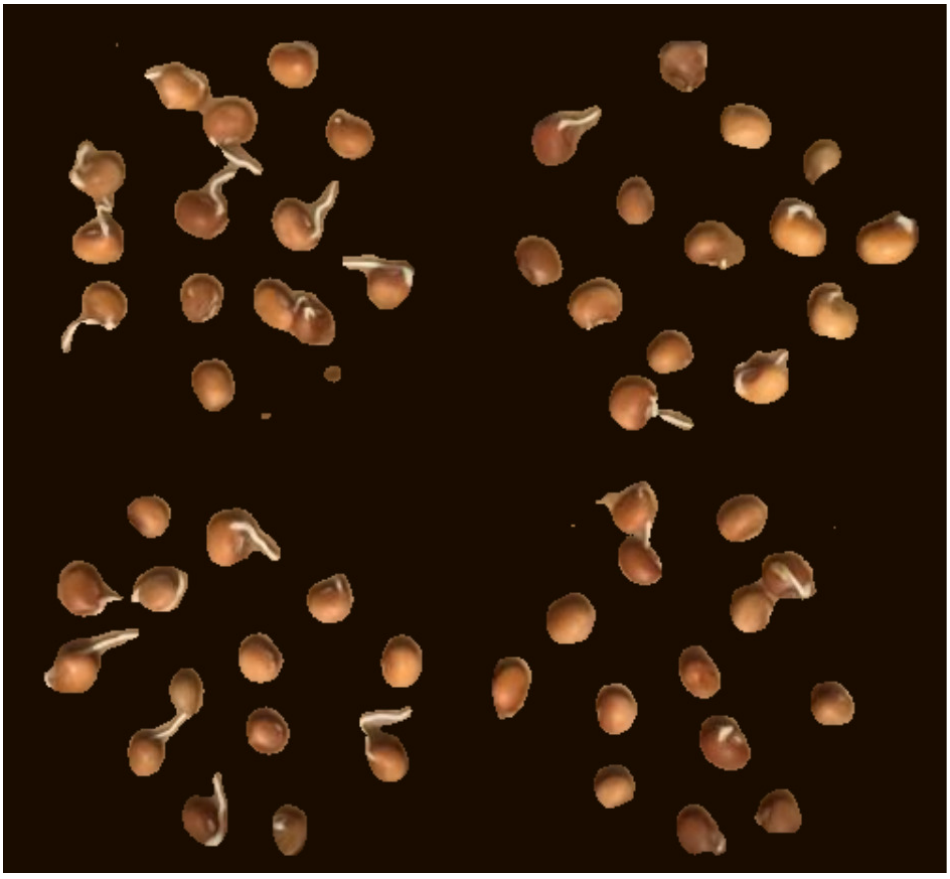


Figure 6.10: Segmented seed images of the proposed SeED-Net system

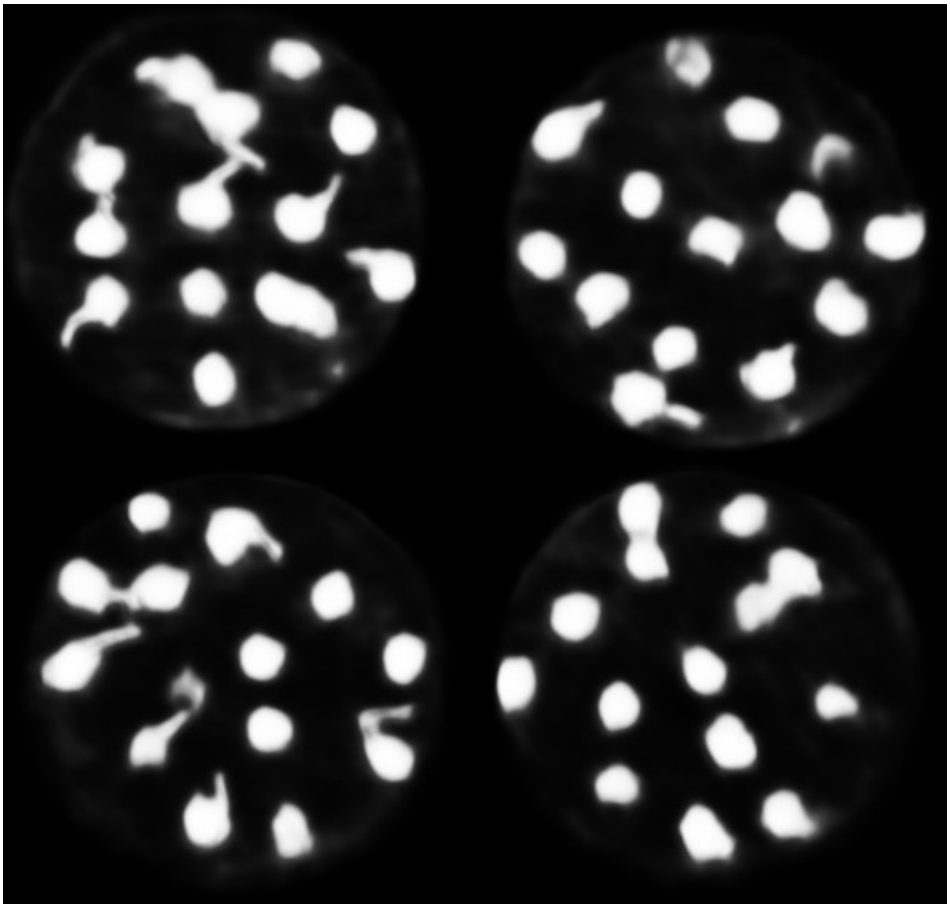


Figure 6.11: Binary image generated from proposed instant segmentation model

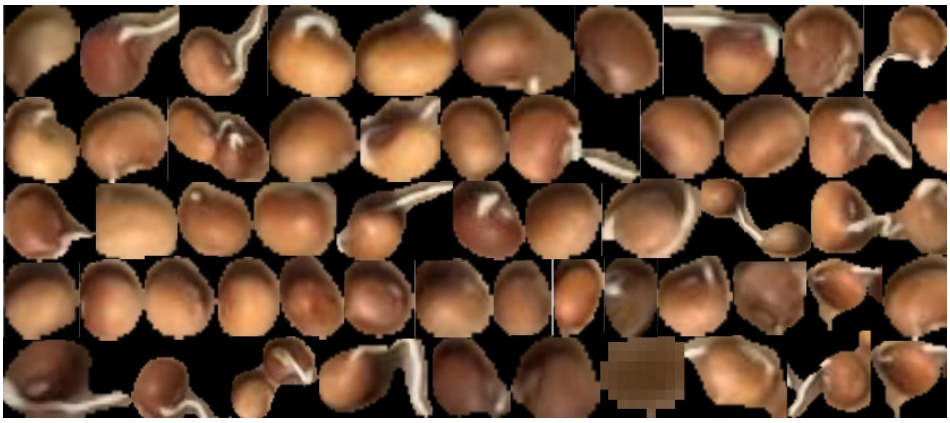


Figure 6.12: Segmented images of the seeds using proposed system

6.7.1 10-Fold Cross Validation (10-FCV)

The 10-FCV is a widely recognised statistical technique used to assess the performance of deep learning models, including the SeED-Ne. This methodology entails dividing the data set into ten sub-samples of roughly equal size. The implementation of SeED-Ne involves the execution of ten iterative cycles for training and validation of the model. In each iteration, a total of nine subsamples are aggregated to create the training data set, while one sub sample is held aside for the purpose of validation. As a result, each individual data point is finally allocated to both the training and validation sets during the iterative process. The overall performance of SeED-Ne is assessed by calculating the average of the individual performance measures after completing all ten cycles as shown in table 6.4. The cyclic structure of this method allows a thorough evaluation of the model, mitigating biases and anomalies. As a result, the SeED-Ne is not only fine tuned but also adaptable and efficient to different seed datasets.Figure 6.13 shows the loss curve from our 10-Fold Cross Validation, a method we used to test the model's accuracy and reliability. By dividing our data into ten parts and testing the model ten times, each time with a diMerent part as the test set, we could see how well our model performed across diMerent data samples. The curve in Figure 6.13 gives us a clear picture of the model's performance, indicating consistent and reliable results across all folds

Table 6.4: 10-Fold validation for proposed SeED-Net model

Model	**Pixel Accuracy**	**IoU**	**Precision**	**Recall**	**F1-Score**
1-Fold	0.82	0.70	0.83	0.83	0.83
2-Fold	0.78	0.76	0.97	0.87	0.85
3-Fold	0.95	0.84	0.88	0.91	0.92
4-Fold	0.96	0.89	0.92	0.94	0.94
5-Fold	0.95	0.89	0.93	0.94	0.94
6-Fold	1.00	0.98	0.98	0.99	0.99
7-Fold	0.98	0.98	1.00	0.99	0.99
8-Fold	1.00	0.97	0.97	0.98	0.98
9-Fold	0.97	0.95	0.98	0.98	0.98
10-Fold	1.00	1.00	1.00	1.00	1.00
Avg	**0.94**	**0.90**	**0.95**	**0.94**	**0.94**

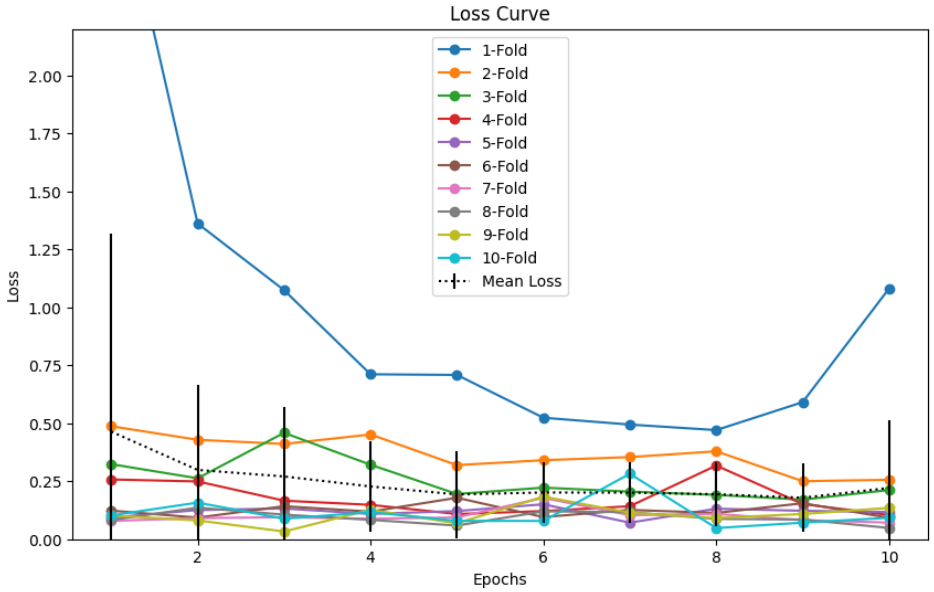


Figure 6.13: 10-Fold Cross Validation loss curve for different models

6.7.2 Optimizer

During the developmental stage of the SeED-Net, several optimizers were tested to evaluate their efficiency in training the network. The aim of the study was to determine the optimizer that would result in the best accuracy, hence ensuring the robust performance of the model in detecting seed and classifying germination. As shown in table 6.5 total of four prominent optimisation algorithms were subjected to rigorous evaluation. Adam, which is an abbreviation for Adaptive Moment Estimation, is a widely utilised optimisation algorithm that calculates adaptive learning rates for individual parameters. Within the framework of SeED-Net, it was able to get a precision rate of 88.07%. The Delta optimizer achieved a level of accuracy of 70.29%. While SeED-Net may be considered a dependable choice in certain situations, it exhibited a comparative delay in comparison to its alternatives in this particular circumstance. Nadam, short for Nesterov-accelerated Adaptive Moment Estimation, can be understood as a hybrid approach that combines the principles of Adam and Nesterov accelerated gradient methods. This approach combines the advantages of the Adam optimizer with the Nesterov's lookahead technique. SeED-Net achieved an accuracy rate of 77.52% in its performance evaluation. The Root Mean Square Propagation optimizer (RMSprop), also referred to as

RMSprop, modifies the Adagrad algorithm by mitigating its overly aggressive and consistently diminishing learning rate. It is worth mentioning that in the experiments carried out using SeED-Net, RMSprop emerged as the highest performing method, exhibiting a noteworthy accuracy rate of 94.00%. Based on the assessments conducted, it can be observed that although all optimizers exhibited their own strengths in specific areas, RMSprop exhibited a notable superiority in terms of accuracy. The qualities of RMS render it a desirable option for integration into the SeED-Net framework, hence guaranteeing the attainment of optimal outcomes in practical seed classification assignments.

Table 6.5: Optimiser's utilized in proposed SeED-Net model

S.No	Optimizer	Training Accuracy
1	ADAM Delta	70.29
2	NaDAM	77.52
3	ADAM	88.07
4	RMS	94.00

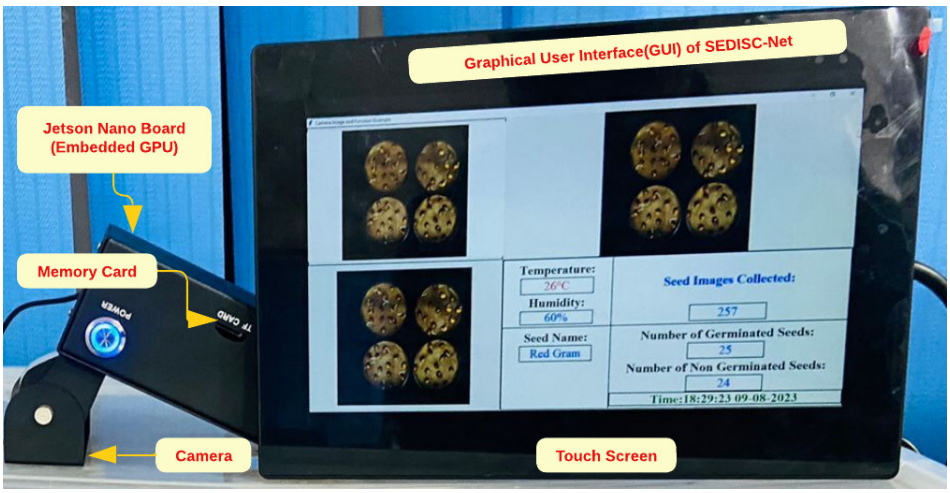


Figure 6.14: Graphical user interface designed for jetson nano for proposed SeED-Net model

6.7.3 Model Deployment on Jetson Nano

The SeED-Ne, a deep learning architecture designed specifically for seed detection and germination classification, was initially trained on an HP Z6 G4 GPU workstation. The comput-

Table 6.6: Evaluation metrics of the fusion models with state-of-art pre-trained models

Model	**Pixel Accuracy**	**IoU**	**Precision**	**Recall**	**F1-score**	**# Parameters**	**Year-Ref**
Detectron + CNN	0.88	0.77	0.89	0.85	0.87	23,02,019	2023 [82]
UNET + ResNet50	0.76	0.65	0.72	0.89	0.79	24,637,826	2023 [83]
UNET + Inception	0.74	0.58	0.77	0.77	0.71	23,903,010	2023 [83]
UNET + CNN	0.91	0.84	0.90	0.92	0.91	23,02,019	2023 [83]
SeED-Net	**0.95**	**0.90**	**0.95**	**0.94**	**0.94**	**3,83,875**	**This Study**

ing capabilities and storage capacity of this high-performance workstation were utilised to efficiently train the model on a substantial dataset, so ensuring the correctness and efficiency of the resultant neural network. Following the completion of the training process, the model underwent optimisation and distillation, resulting in the generation of a final model weight of only 1.8 megabytes. The tiny size of the SeED-Ne architecture offers significant benefits when deployed in situations with limited resources, highlighting its efficiency.

To facilitate real-time prediction of germination in decentralised and on-field scenarios, the SeED-Ne, which is a lightweight model, was implemented on a Jetson Nano embedded GPU. The Jetson Nano is widely recognised for its notable energy efficiency and robust GPU-based processing capabilities. As a result, it has emerged as the optimal platform for executing SeED-Net in situations where prompt processing of first pictures is of utmost importance, even in the absence of extensive computational resources. To enhance user experience and facilitate a full analysis of germination statistics, a specialised Graphical User Interface (GUI) was developed specifically for the deployment of the Jetson Nano as shown in figure 6.14. This user-friendly interface presents essential information to users encompasses; Germination rates refer to a dynamic measure that quantifies the proportion of seeds that have undergone successful germination at a certain point in time. This facilitates the evaluation of seed quality and the efficiency of the germination conditions. Seed Count, the cumulative number of seeds that are processed by the system, providing a concise summary of the sample size under evaluation. The image counter records and analyses the quantity of seed pictures obtained, ensuring a thorough examination of the seed samples. The "Germinated Images Count" refers to a dedicated counter that displays the number of seeds that are identified and categorized as having undergone successful germination by the SeED-Net

system. In contrast this metric indicates the quantity of seeds that have been identified as non-germinated. In summary, the utilization of the efficient and compact SeED-Net on the Jetson Nano embedded GPU not only achieves real-time classification of seed germination in various environments but also offers users a comprehensive and user-friendly representation of germination data, enabling well-informed choices in seed-related agricultural practices.

6.8 Summary

The field of seed detection and germination classification has been significantly impacted by the advent of the Seed Encoding-Decoding Instant Segmentation Network (SeED-Ne), which has emerged as a transformative invention. This model is meticulously designed to effectively address the unique challenges encountered throughout the process of seed germination, exhibiting a remarkable level of precision with an accuracy rate of 94%. The SeED-Ne model demonstrates both accuracy and efficiency when utilized on the high-performance HP Z6 G4 GPU workstation. This is evident from its compact size of merely 1.8mb. The compact nature of this design makes it highly beneficial for implementation on platforms such as the Jetson Nano Embedded GPU, guaranteeing exceptional performance even in limited settings. SeED-Ne, equipped with a user-friendly graphical user interface (GUI) that offers real-time analysis of germination rates, seed counts, and categorization outcomes, presents itself as an essential tool for professionals in the field of agriculture. In the current epoch characterized by increasing global food security demands, the significance of tools such as SeED-Net in augmenting the evaluation of seed quality cannot be overemphasized.

Chapter 7

Conclusion and Future Scope

This chapter presents the summary of the contributions of this thesis, the conclusion of each objective, and the future scope of research is presented.

7.1 Conclusions

In this thesis, a comprehensive exploration of cutting-edge methodologies was undertaken aimed at revolutionizing seed quality assessment and prediction within the realm of agriculture. Through the adept utilization of advanced deep learning techniques, we have endeavored to address the critical challenges faced by modern agricultural practices, thereby fostering the enhancement of crop yields and the promotion of sustainable farming methodologies.

In chapter 3, our research journey involved the deployment of Mask RCNN, a state-of-the-art deep learning architecture, for the purpose of seed quality assessment in conjunction with environmental control measures. By leveraging the capabilities of Mask RCNN that meticulously analyze seed quality attributes while simultaneously accounting for environmental factors that influence seed viability and germination rates. The proposed model architecture demonstrated remarkable efficacy, achieving an accuracy of 84%, as evidenced by the experimental results and the subsequent comparative analysis conducted against other prominent architectures such as ResNet50, Google Inception, and VGG16.

In chapter 4, our investigation delved into the development and implementation of SeedAI,

a novel dual-stage deep learning framework tailored specifically for seed germination prediction. This innovative methodology entailed the utilization of Detectron2 for precise seed extraction followed by classification using a CNN. Through meticulous experimentation, the robustness and reliability of SeedAI is ascertained in accurately predicting seed germination outcomes, achieving an impressive accuracy of 88%, thus underscoring its potential as a valuable tool for agricultural practitioners seeking to optimize planting strategies and maximize crop yields.

In chapter 5, the research endeavors extended to the domain of real-time seed detection and germination analysis, wherein we proposed a fusion model integrating U-Net and CNN architectures deployed on the Jetson Nano platform. This sophisticated framework facilitated semantic segmentation of seeds via U-Net and subsequent germination classification using a custom-designed CNN model. Through rigorous experimentation encompassing dataset collection, data labeling, and comprehensive performance analysis, we validated the efficacy and practical utility of our proposed methodology. Notably, the deployment of the fusion model on Jetson Nano underscored its potential for real-world application in precision agriculture settings, thereby bridging the gap between cutting-edge research and practical implementation, achieving an accuracy of 91%.

In chapter 6, the research culminated in the conceptualization and development of SeED-Net, an innovative seed encoding-decoding network designed to automate seed quality analysis processes. By leveraging instant segmentation techniques and advanced evaluation metrics such as pixel accuracy, Intersection over Union (IoU), precision, recall, and F1 score, SeED-Net showcased unparalleled effectiveness in accurately assessing seed quality attributes, achieving a remarkable accuracy of 94%. The successful deployment of SeED-Net on Jetson Nano further validated its potential for widespread adoption within agricultural contexts, thereby heralding a new era of automation and optimization in seed quality analysis.

In conclusion, the methodologies elucidated in this thesis represent significant advancements in the field of seed quality assessment and prediction within agriculture. By harnessing the power of deep learning and innovative model architectures, we have endeavored to address critical challenges facing modern agricultural practices, thereby paving the way for

enhanced crop yields, sustainable farming methodologies, and a more resilient agricultural ecosystem. Looking ahead, future research endeavors may delve deeper into refining these methodologies, exploring their applicability across diverse agricultural domains, and fostering collaborative efforts to realize the full potential of artificial intelligence in revolutionizing the agricultural landscape. Through sustained innovation and interdisciplinary collaboration, we can aspire to usher in a new era of agricultural productivity, resilience, and sustainability, thereby ensuring food security and prosperity for generations to come.

7.2 Future Scope

As we reflect on the groundbreaking advancements made in seed quality assessment and prediction, it's evident that the journey has only just begun. Looking ahead, there are numerous avenues for further exploration and innovation that hold the potential to reshape the landscape of agriculture in profound ways.

One promising direction for future research lies in refining and enhancing the methodologies we've developed. Hybrid algorithms, optimizing model architectures, and expanding the scope of experimentation may unlock good accuracy and efficiency in seed quality assessment and prediction. This continuous refinement will ensure that our tools remain at the cutting edge of agricultural technology, delivering tangible benefits to farmers and researchers alike. Furthermore, the integration of our methodologies with emerging technologies presents exciting opportunities for innovation. From the Internet of Things (IoT) to blockchain and beyond, there's a wealth of emerging technologies that have the potential to revolutionize agriculture. By leveraging these technologies in conjunction with our deep learning frameworks, we can create synergies that amplify the work and unlock new possibilities for sustainable farming practices.

Bibliography

[1] Wei Long, Shasha Jin, Yujie Lu, Liangquan Jia, Huang Xu, and Linhua Jiang. A review of artificial intelligence methods for seed quality inspection based on spectral imaging and analysis. *Journal of Physics: Conference Series*, 1769(1):012013, 2021.

[2] Costa Caroline, Jácome, Meneghello Geri, Eduardo, Amici Jorge Marçal, Henrique, and Costa Edilson. The importance of physiological quality of seeds for agriculture. *Crop Breeding and Applied Biotechnology*, 17(4):A452, 2021.

[3] Xinyu Zhang. A review of the establishment of the seed lot system in the production of biological products and its importance. *Archives of Razi Institute*, 77, 2022.

[4] María del Carmen Martínez-Ballesta, Catalina Egea-Gilabert, E. Conesa, J. Ochoa, M.J. Vicente, José Franco, Sebastián Bañón, Juan J. Martínez, and Juan A. Fernández. The importance of ion homeostasis and nutrient status in seed development and germination. *Agronomy*, 10:504, 2020.

[5] Conrado Dueñas Adriano Griffo Shraddha Shridhar Gaonkar Francesca Messina Alma Balestrazzi Macovei Anca Andrea Pagano, Paola Pagano. Seed quality assessment and improvement between advancing agriculture and changing environments. pages 317–334, 2023.

[6] B. S. Olisa. Editorial: Seed science and technology. *Seed Science and Technology*, 51:137–143, 2023.

[7] C. Vanitha Devendra K. Yadava S. Sundareswaran, P. Ray Choudhury. Seed quality: Variety development to planting—an overview. pages 1–16, 2023.

[8] B. Badiru Adedeji. Testing seed for quality. pages 299–334, 2023.

[9] Olha Samoilichenko and V.M. Mokiichuk. The evaluation of the measurement uncertainty of the thousand-seed weight in accredited testing laboratories. *Ukraïnskij metrologìčnij žurnal*, 2023.

[10] Edgar Magdaleno Hernández, Adalberto Magdaleno Hernández, Apolinar Mejía Contreras, Tomás Martínez Saldaña, Mercedes A. Jiménez Velázquez, Julio Sánchez Escudero, and José Luis García Cué. Evaluación de la calidad física y fisiológica de semilla de maíz nativo. *Agricultura, Sociedad y Desarrollo*, 17:569–581, 2020.

[11] Sandeep Nagar, Prateek Pani, Raj Nair, and Girish Varma. Automated seed quality testing system using gan active learning. *arXiv: Computer Vision and Pattern Recognition*, 2021.

[12] Eliya Suita and Dida Syamsuwida. Physical characteristics and germination testing methods of turi (sesbania grandiflora (l.) pers) seeds. 5:125–135, 2017.

[13] Igor Djurdjic, Branka Govedarica, and Vesna Milic. Testing living capabilities of seeds of wheat varietys from banja luka (cold test). *Agriculture and Forestry*, 62, 2016.

[14] B. Ramakanth Reddy B. Sai Rushwanth B. Vamsi T. Dhiliphan Rajkumar, P. Nagaraj and G. Dinesh Kumar. Analysis of seed testing to improve cultivation using image processing techniques. pages 804–808, 2023.

[15] Júlia Martins Soares, André Dantas de Medeiros, Daniel Teixeira Pinheiro, Jorge Tadeu Fim Rosas, Daniel Lucas Magalhães Machado, and Denise Cunha Fernandes dos Santos Dias. Low-cost system for multispectral image acquisition and its applicability to analysis of the physiological potential of soybean seeds. *Acta Scientiarum-Agronomy*, 45, 2022.

[16] Chongyuan Zhang, Yongsheng Si, Jacob Lamkey, Rick A. Boydston, Kimberly A. Garland-Campbell, and Sindhuja Sankaran. High-throughput phenotyping of seed/seedling evaluation using digital imaging. *Agronomy*, 8:63, 2018.

[17] Deep learning network of amomum villosum quality classification and origin identification based on x-ray technology. *Foods*, 2023.

[18] Determination of rice seed vigor by low-field nuclear magnetic resonance coupled with machine learning. *INMATEH-Agricultural Engineering*, 2022.

[19] Smart farming using deep learning. *International Journal of Scientific Research in Science and Technology*, 10:371–378, 2023.

[20] Application of plant phenotype extraction using virtual data with deep learning. *Journal of Physics: Conference Series*, 2356:371–378, 2022.

[21] M.R. Nithya, P. Lakshmi, J. Roshmi, R. Sabana, and R.Udhaya Swetha. Machine learning and iot based seed suggestion: To increase agriculture harvesting and development. pages 1202–1207, 2023.

[22] S. Visalakshi and V. Radha. An enhanced k-means clustering based outlier detection techniques to improve water contamination detection and classification. pages 303–313, 2015.

[23] Worasit Sangjan, Nisit Pukrongta, Arron H Carter, and et al. Development of iot-based camera system for automated in-field monitoring to support crop breeding programs. *Authorea*, pages 1–6, 2022.

[24] B. Q. Chen, Z. H. Wu, and H. Y. Li. Research progress in agricultural application of machine vision technology. *Science and Technology Review*, 36(11):54–65, 2018.

[25] A. Chaugule and N. Suresh. Evaluation of texture and shape features for classification of four paddy varieties. *Journal of Engineering*, 2014:1–8, 2014.

[26] M. Abtahi, A. Mirlohi, N. Sharif-Moghaddam, and E. Ataii. Revealing seed color variation and their possible association with yield and quality traits in a diversity panel of flax (linum usitatissimum l.). *Frontiers in Plant Science*, 13:1038079, 2022.

[27] W. J. Elisens and A. S. Tomb. Seed morphology in new world antirrhineae (scrophulariaceae): systematic and phylogenetic implications. *Plant Systematics and Evolution*, 142:23–47, 1983.

[28] B. Lurstwut and C. Pornpanomchai. Image analysis based on color, shape, and texture for rice seed (oryza sativa l.) germination evaluation. *Agriculture and Natural Resources*, 51:383–389, 2017.

[29] P. Khoenkaw. An image-processing based algorithm for rice seed germination rate evaluation. In *Proceedings of the 2016 International Computer Science and Engineering Conference (ICSEC)*, page 1–5, Chiang Mai, Thailand, 2016. IEEE.

[30] S. Ducournau, A. Feutry, P. Plainchault, P. Revollon, B. Vigouroux, and M. Wagner. An image acquisition system for automated monitoring of the germination rate of sunflower seeds. *Computers and Electronics in Agriculture*, 44:189–202, 2004.

[31] Agripheno, lemnatec launched new seed germination detection system germination scanalyzer, 2022.

[32] Sc-h mobile phone photo seed automatic counting machine, 2022.

[33] Multi-feature machine learning model for automatic segmentation of green fractional vegetation cover for high-throughput field phenotyping. *Plant Methods*, 2017.

[34] S. Barmpounakis, A. Kaloxylos, A. Groumas, L. Katsikas, V. Sarris, K. Dimtsa, F. Fournier, E. Antoniou, N. Alonistioti, and S. Wolfert. Management and control applications in agriculture domain via a future internet business-to-business platform. *Information Processing in Agriculture*, 2:51–63, 2015.

[35] K. Jha, A. Doshi, P. Patel, and M. Shah. A comprehensive review on automation in agriculture using artificial intelligence. *Artificial Intelligence in Agriculture*, 2:1–12, 2019.

[36] V. G. Dhanya, A. Subeesh, N. L. Kushwaha, D. K. Vishwakarma, T. Nagesh Kumar, G. Ritika, and A. N. Singh. Deep learning based computer vision approaches for smart agricultural applications. *Artificial Intelligence in Agriculture*, 6:211–229, 2022.

[37] H. Tian, T. Wang, Y. Liu, X. Qiao, and Y. Li. Computer vision technology in agricultural automation—a review. *Information Processing in Agriculture*, 7:1–19, 2020.

[38] N. Lakshmi Kalyani and Kolla Bhanu Prakash. Soil color as a measurement for estimation of fertility using deep learning techniques. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 13:305, 2022.

[39] Predicting the quality of soybean seeds stored in different environments and packaging using machine learning. *Dental science reports*, 2022.

[40] Seedquant: A deep learning-based tool for assessing stimulant and inhibitor activity on root parasitic seeds. *Plant Physiology*, 2021.

[41] A prediction method of seedling transplanting time with dcnn-lstm based on the attention mechanism. *Agronomy*, 2022.

[42] Quantitative evaluation of maize emergence using uav imagery and deep learning. *Remote sensing*, 2023.

[43] Generative models-based data labeling for deep networks regression: application to seed maturity estimation from uav multispectral images. *Remote sensing*, 2022.

[44] Predicting plant growth from time-series data using deep learning. *Remote Sensing*, 2021.

[45] Robust seed germination prediction using deep learning and rgb image data. *Scientific Reports*, 2021.

[46] Comparative analysis of different artificial neural networks for predicting and optimizing in vitro seed germination and sterilization of petunia. *PLOS ONE*, 2023.

[47] John Heslop-Harrison. Plant development. 2022.

[48] V. G. Dhanya et al. Deep learning-based computer vision approaches for smart agricultural applications. *Artificial Intelligence in Agriculture*, 6:211–229, 2022.

[49] R. P. S. Testing seed for quality. 2023.

[50] Q. Peng et al. Automatic monitoring system for seed germination test based on deep learning. *Journal of Electrical and Computer Engineering*, 2022:4678316, 2022.

[51] M. L. Srivastava. Seed germination, mobilization of food reserves, and seed dormancy. In *Plant Growth and Development*, page 447–471. Academic Press, San Diego, CA, USA, 2002.

[52] H. Khaeim et al. Impact of temperature and water on seed germination and seedling growth of maize i (zea mays l.). *Agronomy*, 12(2):397, 2022.

[53] R. Anju et al. Developing climate-resilient chickpea involving physiological and molecular approaches with a focus on temperature and drought stresses. *Frontiers in Plant Science*, 10:1759, 2020.

[54] A. Kamilaris and F. X. Prenafeta-Boldú. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147:70–90, 2018.

[55] L. Benjamaporn and C. Pornpanomchai. Application of image processing and computer vision on rice seed germination analysis. *International Journal of Applied Engineering Research*, 11:6800–6807, 2016.

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

[57] F. Xijian, Z. Rui, T. Tardi, D. C. Sruti, and Y. Qiaolin. A segmentation-guided deep learning framework for leaf counting. *Frontiers in Plant Science*, 13:844522, 2022.

[58] A. Walter, F. Liebisch, and A. Hund. Plant phenotyping: From bean weighing to image analysis. *Plant Methods*, 11:14, 2015.

[59] Q. Xiao, X. Bai, C. Zhang, and Y. He. Advanced high-throughput plant phenotyping techniques for genome-wide association studies: A review. *Journal of Advanced Research*, 35:215–230, 2022.

[60] F. Forcella, R. L. B. Arnold, R. Sanchez, and C. M. Ghersa. Modeling seedling emergence. *Field Crops Research*, 67(2):123–139, 2000.

[61] Yanchao Zhang, Jiya Yu, Yang Chen, Wen Yang, Wenbo Zhang, and Yong He. Real-time strawberry detection using deep neural networks on embedded system (rtsd-net): An edge ai application. *Computers and Electronics in Agriculture*, 192:106586, 2022.

[62] N. A. George–Jones et al. Automated detection of vestibular schwannoma growth using a two-dimensional u-net convolutional neural network. *Laryngoscope*, 131(2):E619–E624, 2021.

[63] K. Hadi et al. Development of pixel-wise u-net model to assess performance of cereal sowing. *Biosystems Engineering*, 208:260–271, 2021.

[64] C. Qian et al. An improved u-net network-based quantitative analysis of melon fruit phenotypic characteristics. *Journal of Food Measurement and Characterization*, 16(5):4198–4207, 2022.

[65] W. Choi and Y.-J. Cha. Sddnet: Real-time crack segmentation. *IEEE Transactions on Industrial Electronics*, 67(9):8016–8025, 2020.

[66] R. Ali and Y.-J. Cha. Attention-based generative adversarial network with internal damage segmentation using thermography. *Automation in Construction*, 141:104412, 2022.

[67] J. Lewis, Y. J. Cha, and J. Kim. Dual encoder-decoder-based deep polyp segmentation network for colonoscopy images. *Scientific Reports*, 13(1):1183, 2023.

[68] K. Cao and X. Zhang. An improved res-convolutional model for tree species classification using airborne high-resolution images. *Remote Sensing*, 12(7):1128, 2020.

[69] K. V. Suma, D. B. Koppad, K. Awasthi, A. K. Phani, and R. Vikas. Application of ai models in agriculture. *Proceedings of the 4th International Conference on Circuits, Control, Communication and Computing (I4C)*, pages 387–390, 2022.

[70] Luohan Wang, Diming Zhang, and Zhenxing Li. Clothing pattern recognition based on attention mechanism and mask rcnn network. 12721:1272111, June 2023.

[71] Prakash S R and Nath Singh P. Object detection through region proposal based techniques. *Materials Today: Proceedings*, 46(Part 9):3997–4002, 2021.

[72] Roboflow website. https://roboflow.com/.

[73] G Damour, C Guérin, and M Dorel. Dataset on early growth of cover crops in growth chamber. *Data in Brief*, 29:105262, 2020.

[74] Introduction to raspberry pi. 2019.

[75] Roboflow: A flow-based visual programming language for mobile manipulation tasks. 2015.

[76] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *Computers in Biology and Medicine*, 81:79–92, 2017.

[77] A novel unet segmentation method based on deep learning for preferential flow in soil. *Soil Tillage Research*, 2023.

[78] Ruyue Xin, Jiang Zhang, and Yitong Shao. Complex network classification with convolutional neural network. *Tsinghua Science and Technology*, 25(4):447–457, August 2020.

[79] Characterizing the performance of accelerated jetson edge devices for training deep learning models. 2023.

[80] Pi usb cam: A simple and affordable diy solution that enables high-quality, high-throughput video capture for behavioral neuroscience research. *ENeuro*, 2022.

[81] Dmitry Kolosov, Vasileios Kelefouras, Panagiotis Kourtessis, and Iosif Mporas. Anatomy of deep learning image classification and object detection on commercial edge devices: A case study on face mask detection. *IEEE Access*, 10:109167–109186, 2022.

[82] Prakash Kodali Ramesh Reddy D, RamalingaSwamu Cheruku. Seedai: a novel seed germination predictionsystem using dual stage deep learning framework. *Environmental Research Communications*, 5:125014, 2023.

[83] R. R. Donapati, R. Cheruku, and P. Kodali. Real-time seed detection and germination analysis in precision agriculture: A fusion model with u-net and cnn on jetson nano. *IEEE Transactions on AgriFood Electronics*, 1(2):145–155, Dec 2023.

List of Publications

Publications from the thesis

Patent:

1. **Ramesh Reddy Donapati,**, R. Cheruku and Prakash Kodali, “SALT TOLERANCE MONITORING SYSTEM” at Germination stages of seeds”, Patent Number: 454893, **(Granted), 2023**.

Journal papers:

1. **Ramesh Reddy Donapati,**, R. Cheruku and Prakash Kodali, “SeedAI: a novel seed germination predictionsystem using dual stage deep learning framework” ***Environmental Research Communications***, Volume 5, Number 12, 125014, 2023. **(Published)**.
   DOI 10.1007/s00500-023-08393-5
2. **Ramesh Reddy Donapati**, R. Cheruku and Prakash Kodali, "Real-Time Seed Detection and Germination Analysis in Precision Agriculture: A Fusion Model With U-Net and CNN on Jetson Nano" ***IEEE Transactions on AgriFood Electronics***, Volume 1, no.2, pp. 145-155,Dec 2023. **(Published)**.
   DOI: 10.1109/TAFE.2023.3332495.
3. **Ramesh Reddy Donapati**, R. Cheruku and Prakash Kodali "SeED-Net: Seed Encoding Decoding Network for Enhancing Seed Quality Analysis" ***IEEE Transactions on AgriFood Electronics*** - **(Under Review)**.

4. **Ramesh Reddy Donapati**, R. Cheruku and Prakash Kodali "AI-Enhanced Seed Quality Assessment and Environmental Control Using MASK RCNN” ***"Geospatial Artificial Intelligence (GeoAI) Applications in Agriculture for Smart Farming Solutions"***, 2023.– **(Under Review)**

Conference papers:

1. **D. Ramesh Reddy**, K. Naga Santhosh and P. Kodali, "Convolutional Neural Networks for the Intuitive Identification of Plant Diseases," 2022 International Conference on Inventive Computation Technologies (ICICT), Nepal, 2022, pp. 1-6, doi: 10.1109/ICICT54344.2022.9850695, **(Published)**

2. **D. Ramesh Reddy**, M. Kavya, S. Dharani, S. S. Tumpudi, P. Kodali and N. Sandhya, "Design and Development of a Low-Cost Crop Protection System Using the Internet of Things and Machine Learning," 2022 IEEE International Symposium on Smart Electronic Systems (iSES), Warangal, India, 2022, pp. 610-614, doi: 10.1109/iSES54909.2022.00134. **(Published)**