Deep Learning-based Optical Technique for Early Identification and Detection of Seed-Borne Disease

M.Tech. Thesis

By

Gawali Siddhesh Rajendra Savita

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

JUNE 2022

Deep Learning-based Optical Technique for Early Identification and Detection of Seed-Borne Disease

A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree

of

Master of Technology

by

Gawali Siddhesh Rajendra Savita

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

JUNE 2022

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **Deep Learning-based Optical Technique for Early Identification and Detection of Seed-Borne Disease** in the partial fulfilment of the requirements for the award of the degree of **MASTER OF TECHNOLOGY** and submitted in the **DEPARTMENT OF ELECTRICAL ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from August 2020 to June 2022 under the supervision of **Dr. Vimal Bhatia, Professor, IIT Indore**

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

31-05-2022

Signature of the student with date
(Gawali Siddhesh)

This is to certify that the above statement made by the candidate is correct to the best of my/our knowledge.

31.05.2022

Signature of the Supervisor of M.Tech. thesis (with date)
(Prof. Vimal Bhatia)

Gawali Siddhesh Rajendra Savita has successfully given his/her M.Tech. Oral Examination held on **06/06/2022**.

06.06.2022

Signature(s) of Supervisor(s) of M.Tech. thesis
Date:

Convener, DPGC
Date: 06/06/2022

06.06.2022

Signature of PSPC Member #1
Date:

Signature of PSPC Member #2
Date: 06-06-2022

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis supervisor Dr. Vimal Bhatia for the constant support, motivation, and guidance throughout my project work. I would also like to thank Dr. Shashi Prakash for his constant support and guidance.

I would like to express my gratitude to my PSPC committee member Dr. Vivek Kanhangad, Dr. Nemnath Hubballi, for their valuable suggestions and insightful comments. Their questions and suggestions helped me to widen my knowledge from various perspectives.

I would also like to thank the Department of Electrical Engineering, IIT Indore for its support and for providing the required lab facilities. I would also like to the IISR Indore for providing seed samples.

I would also like to thank all the faculty members of the Electrical Engineering Department for teaching me during my course work which led a to strong foundation for this work.

I would like to thank my mentor Ph.D. scholar Puneet Singh Thakur for constant motivation and encouraging me throughout my project work.

I am especially grateful to my friends Bhagyashree, Raj, Shubham and Utkarsha for being there during my hardships.

Lastly, I am greatly thankful to my family for their true support in pursuing my goals.

Siddhesh Gawali

M.Tech (Communication and Signal Processing)

2002102012

Department of Electrical Engineering, IIT Indore

Dedicated

to my

La Petite Sister

Abstract

The agricultural sector is the backbone of the country's economy. There are a variety of factors that affect crop production, with seed-borne diseases being one of the most important. Seeds and grains are the agricultural sector's backbone, making them critical assets to manage and preserve both pre- and post-harvest. There are different optical techniques for the analysis and processing of diseased seed samples employing various image acquisitions such as laser biospeckle, infrared imaging, and so on, as shown in the literature research. However, traditional image processing systems have some processing and experimentation limitations. As a result, a deep learning-based optical technique processing pipeline for detecting and identifying diseased and healthy seed samples using laser speckle patterns in this study.

Transfer learning and ensemble learning-based algorithms and models are used in the proposed study. AlexNet, VGG16, ResNet18, GoogleNet, and MobileNetV2 are the best five state-of-the-art transfer learning-based CNN models, with ResNet18 having the greatest accuracy of 95.33%. Along with TL models, a CNN model with an accuracy of 95.5 % was trained on the target dataset as a baseline model for comparing it to other models. Finally, the (re-)trained models are used in an ensemble learning-based strategy. The algorithms Majority Vote and Early Fusion are used. The accuracy of the Majority Vote model, which included all five (re-)trained models, was calculated to be 96.1 %. The accuracy was calculated as 98.1% for the early fusion model, which incorporated the top three models with the best individual test accuracies. The results demonstrate the use of ensemble learning models exceeds the performance of individual state-of-the-art models for the targeted dataset and desired task.

LIST OF PUBLICATIONS

Under Review:

Puneet Singh, Siddhesh Gawali, Vimal Bhatia and Shashi Prakash, "Deep learning-based processing framework for spatio-temporal analysis and classification of laser biospeckle data," Optics and Lasers in Engineering.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AVD	Absolute Value Difference
BCE	Binary Cross-Entropy
CCD	Charge-Couple Device
CNN	Convolutional Neural Network
COM	Co-Occurrence Matrix
Conv	Convolution Operation
DL	Deep Learning
DNN	Deep Neural Network
EF	Early Fusion
FC	Fully Connected
FN	False Negative
FP	False Positive
FPS	Frames Per Second
GD	General Difference
GDP	Gross Domestic Product
He-Ne	Helium-Neon
IM	Inertia Moment
MHI	Motion History Image
MV	Majority Vote
MXP	Max-Pooling Operation
ReLU	Rectified Linear Unit
RGB	Red Green Blue
RTHSP	Random Temporal History of Speckle Pattern
SVM	Support Vector Machine
THSP	Temporal History of Speckle Pattern
TL	Transfer Learning
TN	True Negative
TP	True Positive

Chapter 1

Introduction

According a recent survey, India's economy major depends on agricultural sector which contributes 20.19% of GDP's share (Ministry of Statistics and Programme Implementation, 2021). There are multiple factors that have impact on crop production out of which seed-borne diseases is an important factor. The seeds and grains are the backbone of agricultural sector, making them the essential assets to maintain and preserve pre-and post-harvest. According to the study (Rajput et al., 2020), a few contagious diseased seeds can ruin the entire harvest or the stored batch. This motivates us to explore and understand the challenges faced by the agricultural sector and make an attempt to resolve the issue.

It is crucial and necessary to identify the diseased seed before storing and shipment process of the harvest. Various factors can be responsible for seed-borne diseases, such as pathogen attack, inhabitable moisture levels, unusual weather conditions, and many others. However, traditional physical techniques exist, which include chemical and physical techniques to analyze seeds under observation. But this approach has several drawbacks, such as the involvement of chemical processes, the need for seed to be at a specific germination stage, the need for highly skilled manpower and most important these techniques can be invasive or destructive in nature.

But in the past several successful techniques which are non-invasive in nature such as Infrared imaging (Liu et al., 2019), X-ray imaging (Sood et al., 2016), laser biospeckle imaging (Cardoso et al., 2011), and others. These strategies have been demonstrated to be reliable in a variety of sectors, including seed quality detection. However, these techniques have their own drawbacks and with respect to hardware setup complexity, equipment mobility, required for a controlled environment for imaging, need for high complexity analytical techniques.

Laser biospeckle imaging (Zdunek et al., 2014) has improved over the past few decades with the involvement of complex optical processing techniques and technological advancements in image acquisition hardware setup. However, the techniques proposed in the literature possess several experimental- and processing-related drawbacks which provides manual visual features for processing. In principle, these techniques-controlled input with specific image quality parameters such as frame size, number of frames, image resolution, etc. This encourages us to explore the application and deployment of DL-based methods for the analysis and processing of seed samples for seed health inspection.

Hence, to overcome the drawbacks od discussed, in this thesis we propose a DL-based optical technique for early detection and classification for seed-borne disease. seeds.

Organization of thesis

This chapter provides a brief overview of the difficulties faced in the agricultural sector and how they are addressed, laying the groundwork for our desire to use laser backscattering imaging to detect healthy and disease seeds. The thesis' remaining contents are arranged as follows:

- Chapter 2: This chapter includes the basic theory required to understand laser backscattering patterns and the hardware setup required for laser backscattering image acquisition.
- Chapter 3: This chapter includes a brief literature review of the prosed techniques proposed and used in the past for processing and analysis of laser backscattering images.
- Chapter 4: In this chapter, a summarised theory about deep learning and related topics will help in understanding the further section.
- Chapter 5: The suggested DL-based processing and analysis framework will be discussed in this chapter. This chapter will include transfer learning models and an ensemble model, as well

as an ensemble learning-based model that is offered as a better alternative to the traditional DL method.

- Chapter 6: A summary of the target dataset, as well as a brief overview of the work's findings, are included in this chapter. After evaluating numerous models, the best design is suggested based on accuracy and other criteria.
- Chapter 7: The proposed work's conclusion, as well as possible future scope for expanding this work to additional applications and domains.

Chapter 2

Laser Backscattering Technique

This chapter includes the basic theory required to understand laser backscattering patterns and the hardware setup required for laser backscattering image acquisition.

2.1 Speckle Pattern Imaging

Speckle images (Zdunek et al., 2014) are generated due to biological material being illuminated with coherent light or laser light. This phenomenon results from laser backscattering which gives rise to a static interference pattern (Ibrahim, 2016). When the light illuminates the biological surface, the laser penetrates a few mm into the surface and is reflected back. Inside the biological object, the light undergoes absorption and scattering, causing a speckle pattern to appear on the detecting plane. This speckle can further be used to analyse multiple physiological and biological factors. The bright and dark pixelated regions, as shown in Fig. 1, result from mutual self-interference between backscattering light having the same frequency but different phases. As biospeckle is a non-destructive method of extracting biological activity information, and due to penetration phenomenon has proven more informative than simple RGB imaging (Kumari & Nirala, 2019).

Moreover, this technique possesses a considerable number of advantages, such as robustness to vibration and external noise a more straightforward process for acquiring images, and a basic processing pipeline that takes lesser memory requirements and is compatible with various image processing and processing algorithms. When a single speckle is captured, the pattern is static in nature. When the object under study is inanimate, the speckle pattern is temporally static. First, light enters the tissue and can be backscattered from either the surface or the internal inhomogeneity beneath the tissue in the case of living organisms. Secondly, if the object is living, neither the chemical nor the physical processes are stalled in the sample. This leads to speckle patterns having time-varying properties. This is generally referred to as a dynamic speckle pattern. In (Draijer et al., 2008), the author suggests that this dynamic behaviour can be due to the motion of the particle inside the sample and doppler shift.

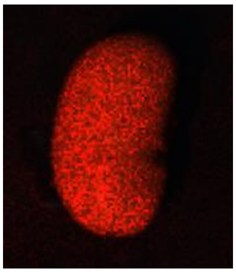


Figure 1 Laser backscattering pattern of a diseased soybean seed

Depending on the light scattering in the material under examination, the intensity shift in the speckle pattern is very unpredictable. As a result, the speckle pattern collected has a biological signature of the phenomena occurring within the sample. The sum of numerous components with a phase shift of Φ_n and an amplitude of a_k dispersed from diverse places on the recorded field is the result of light intensity caught at an arbitrary point F(x,y). The speckle frame created by the surface element j at position F may be mathematically represented as (Kalyzhner et al., 2019):

$$\boldsymbol{I_j(F)} = \sum_{\boldsymbol{n=1}}^{\boldsymbol{N}} \boldsymbol{a_k} \mathrm{e}^{(\mathrm{i}\boldsymbol{\delta\phi_n})} \quad (1)$$

where, (iota), $i = \sqrt{(-1)}$ and N is the total number of secondary wave pattern created. These recorded fluctuations in intensity, as well as a variety of other speckle characteristics (such as speckle grain size, spatial and temporal fluctuation of pixels, contrast, roughness, and so on.), are exceedingly difficult to anticipate using traditional statistics based on image processing techniques.

Due to its non-invasive and non-contact property, this imaging techniques finds its application in many fields of medical, agriculture, spoof detection and many more. Some of its major and well know applications are; plant disease detection (Dhanotia et al., 2017), static scatterer concentration assessment in phantom bodily fluids (Jayanthy et al., 2011), fruit bruise assessment (Pajuelo et al., 2003), biometric identification spoof detection (Chatterjee et al., 2017, 2018, 2019), priming analysis (Singh et al., 2020), imaging of the parathyroid and cancer cells (Braga et al., 2012)and many more.

2.2 Experimental Hardware Setup

The above technique is used for data acquisition of dataset images. This requires a specific hardware setup comprising of He-Ne laser, spatial filter, magnifier, CCD Camera, and a vibration isolation tabletop with sample space. Figure 2 depicts the hardware stack. A variable attenuator is then used to regulate the laser's intensity. The laser beam is expanded using a spatial configuration consisting of a microscopic object (MO) with a total magnification of 10 m and an aperture of 10 m. This filtering approach also reduces the non-uniformity in the laser profile produced by noise, resulting in an evenly lighted laser-illuminated zone. The sample under investigation is illuminated with a He-Ne diode laser (λ= 632.8 nm, 15 mW). The seed is then placed on the vibration isolation table in the region illuminated by the laser. A charged-coupled device camera (Basler Corp., resolution 1024 x 967, frame rate: 32 fps) records the speckle patterns that correspond to each seed sample. Later this captured data is processed, and the dataset is obtained with desired quality parameters like saturation, contrast, homogeneity, and others in MATLAB.

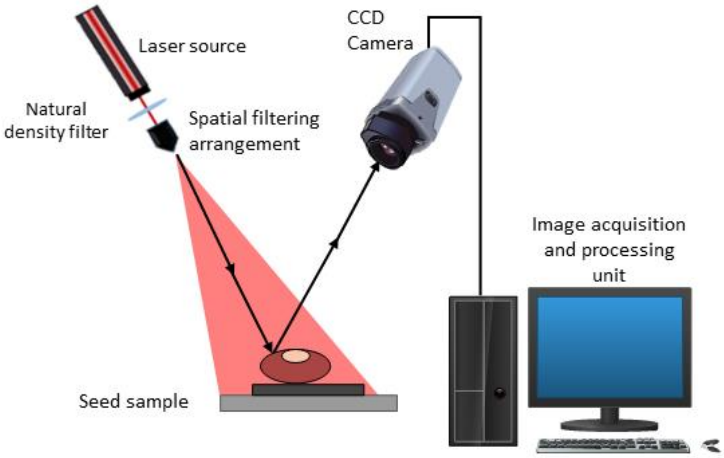


Figure 2 Experimental hardware setup for recording speckle images

.

Chapter 3

Literature survey

This chapter includes a brief literature review of the prosed techniques proposed and used in past for processing and analysis of laser backscattering images.

For the assessment and analysis of the laser backscattering data, various image processing techniques have been proposed in the literature. The strategies rely heavily on mathematical modeling of the processing pipeline for assessing biological activity and have demonstrated a variety of ways, each with its own set of advantages and disadvantages. At the current age, very few of these conventional approaches comprising significant use of manual image processing algorithms have held up to the needs of the modern age's real-time, low latency, low computational processing needs.

Furthermore, these strategies may be classified into numerical and visual categories. The numerical techniques (Oulamara et al., 1989) comprise a framework with a numerical value or an index as an output. This index is used as a measurement to identify the sample's biological activity on the scale provided by the respective numerical technique. On the other hand, the visual techniques (Fujii et al., 1985) have a feature map indicating the sample's portion having the highest or most recent biological activity. These techniques are used to identify the region of the sample having disease or abnormalities (Jayanthy et al., 2011).

3.1 Numerical Techniques

The Temporal History of Speckle Pattern or Random THSP, together with Inertia Moment and Co-Occurrence Matrix, is a regularly used numerical indexing approach strategy (Alves et al., 2013; Dhanotia et al., 2017; Pajuelo et al., 2003; Zdunek et al., 2014). The THSP (Zdunek et al., 2014) is generated by extracting a single from every speckle frame from the same row or column position and placing it side by side. Whereas, in RTHSP. A random row or column is taken from each frame instead of from the same position from all frames. The width can vary up to the number of available speckle frames.

The sample functions as the range of non-zero values other than the primary diagonal values, and the COM is derived using the THSP matrix. The sample is likely to be less active if the values in COM are concentrated towards the primary diagonal, whereas a sample with a wide range of values in COM's diagonal is expected to be more active. To quantify the activities from COM, IM is defined as the sum of the matrix values times the squared distance from the principal diagonal. Higher activity equates to a higher IM value, while lower activity corresponds to a lower IM value.

However, according to (Arizaga et al., 1999; Braga et al., 2012), it has been seen that IM shows higher sensitivity for higher activity or variations and performs poorly for medium and low activity. Therefore (Braga et al., 2011), AVD was proposed instead of IM by taking first-order difference to increase the robustness of the numerical value.

3.2 Visual Techniques

The visual techniques are considered full-field techniques (Briers, 1996), i.e., it utilizes every pixel of each speckle frame for generating a feature map containing information about the region with the highest activity or with the most recent activity in the sample under study. Being full-field techniques, some of the visual techniques have proven to have higher accuracy, zero standard deviation, capability to process both homogeneous and heterogeneous biospeckle patterns.

The first visual technique proposed was Fujji’s Method (Fujii et al., 1985, 1987) for blood flow observation using dynamic speckle patterns. The method uses a weighted sum of the differences between two successive components of each pixel's temporal sequence of intensities. The following equation can be used to calculate the index.

$$P(x,y) = \sum_{n=1}^{N} \frac{|K_n(x_1y) - K_{n+1}(x,y)|}{|K_n(x_1y) + K_{n+1}(x,y)|} \qquad (2)$$

where n is the image index ranging from n = 1 to N and K_n is the n^{th} frame in sequence having (x,y) as pixel coordinates. The image composed of the integrated P(x,y) values having biological activity (high and low). The main disadvantage of Fujji's methods is that they produce a nonlinear response that emphasises both high and low value ranges from the dynamic ranges' boundaries. This sometimes results in addition of noise in the resulting feature map.

The GD (Arizaga, 2002) technique includes the differences between non-consecutive frames as shown in Eqn. 3. This method was proposed as an easier and simpler alternative to Fujji's method to reduce the weighting process.

$$GD(x,y) = \sum_{k=1}^{N-1} \sum_{l=k+1}^{N} |I_k(x,y) - I_l(x,y)| \qquad (3)$$

Because the sequence of appearance is ignored, this method results in a loss of temporal information regarding pixel intensity activity. Furthermore, it merely displays the range of values rather than the frequency of transitions. To overcome this drawback and to reduce the computational time, the absolute operation was replaced by squaring operation as shown in Eqn. 4.

$$GD*(x,y) = \sum_{k=1}^{N-1} \sum_{l=k+1}^{N} (I_k(x,y) - I_l(x,y))^2 \qquad (4)$$

Another technique is to use a motion history image. This is a method of real-time imaging that generates a movement map based on the sample's recent activity. This approach was presented as a replacement for online biospeckle analysis methods (Godinho et al., 2012). First in, first out pipeline design is used to process the images in the buffer. Following the thresholding of the contour, the difference between the two photos was determined. The weighting of the threshold pictures generated in relation to each image's "lifetime" is used to construct the final MHI image. MHI also outperforms offline approaches like Fujji's and GD, according to the study.

3.3 Spatio-Temporal Techniques

Most of the techniques mentioned earlier are considered to be spatial techniques. Spatio-temporal analysis of dynamic speckle patterns using singular value decomposition (Kulkarni et al., 2021) and biospeckle indexing technique employing morphological and geostatistical descriptors are two further ways (Amit Chatterjee et al., 2020). As described by the author, each pattern in the temporal sequence is reduced to a representation using column vectors, which is then aggregated to create a matrix encoded with localized Spatio-temporal speckle intensity data (Kulkarni et al., 2021). The singular values of this matrix are used to establish a correlation metric, which is then used to interpolate the correlation computed at each patch to generate a correlation map. According to (Amit Chatterjee et al., 2020), the N number of the image sequence is captured over some time. Foreground and background separation is done using the following procedures, i.e., lowpass filtering, background texture removal, Otsu's threshold, and binarization. From Otsu's thresholder image, a bounding box defines the intended active section and creates a binary image or mask (for noise reduction of the activity map). With alpha-variogram-based visual analysis, the absolute difference, Eqn. 5, between consecutive images from the image sequence, is calculated to extract temporal variation between consecutive frames. The 'visual variogram estimate' (VVE), $\gamma_v(x,y)$, the function provides a 2D activity map with better performance and a relatively fast operating speed. The mask is multiplied with all the frames one at a time and background noise is suppressed. The output, F(x,y), obtained is used to calculate Numerical indexing using Ensemble averaged dynamicity metric' (EADM), Eqn. 6, where M is the total number of object pixels.

$$\gamma(x) = \frac{1}{2(N-n)} \sum_{k=1}^{N-n} |I_{k+n}(x) - I_k(x)|^{\alpha} \tag{5}$$

$$\hat{a}_{EADM} = \frac{1}{M}\sum_{x=1}^{N_x}\sum_{y=1}^{N_y} F(x, y) \tag{6}$$

However, these above-mentioned techniques used to process biospeckle images have several drawbacks and are dependent on the various experiment- and processing-related factors such as sampling frequency, number frame acquired, illumination condition, region of interest selection, varying background, size of the speckle frames, external noise, and others (Zdunek et al., 2014). Also, these methods consist of manual operations, either the filter a static in nature or require a human to manuals such as labeling region of interest or foreground-background separation, background noise suppression, etc. To eliminate manual operations and human interaction, deep learning can be used to achieve the desired goal in a more optimal way.

Chapter 4

Fundamentals

In this chapter, a summarised theory about deep learning and related topics will help in understanding the further section.

4.1 Deep Learning

In recent years, machine learning has opened new avenues in various fields (Alom et al., 2019). This aids in improving the performance and automizing the workflow requiring minimum human and manual operations. Manual feature extraction based on visual parameters including contrast, mean intensity, weighted difference, correlation index, and others is used in traditional ways for processing speckle pattern frames. Many of the traditional and manual approaches given for the examination and detection of diseased and healthy seeds have been superseded by deep learning algorithms in recent years. (Xuan et al., 2018; Chen et al., 2020; Yao et al., 2021; Yang et al., 2021; Turkoglu et al., 2021). Also, as mentioned in Ch. 3. the proposed techniques in the literature survey have several drawbacks and are dependent on experimental- and processing-related factors such as sampling frequency, number frame acquired, illumination condition, region of interest selection, varying background, size of the speckle frames, external noise and perturbations, and others. Therefore, to overcome these drawbacks, a optical technique based on deep learning algorithms that outperform manual and traditional methods for identifying healthy and diseased seeds.

4.2 Convolutional Neural Network

CNNs (Lecun et al., 2015) are a type of deep learning model that is used for image recognition and classification. The models are trained on specific targeted datasets, for a particular task. The parameters are initialized to random weights when training a CNN from scratch. Fig. 3.

depicts the general architecture of CNN. The convolutional layer is accountable for feature extraction. During the training and validation phases, these characteristics are automatically optimized as the number of epochs rises. The kernel size varies depending on the learner's and task's needs. Convolution operation provides images with edge detection, image blurring, etc. mathematically the feature map can be expressed as (Lecun et al., 2015);

$$x_j^l = g(\sum_{i \in N_j} x_i^{l-1} * k_{ij}^l + b_j^l) \tag{7}$$

where x_j^l is the output of current layer, xil-1 is output of previous layer, k_{ij}^l are the kernel weight of the current layer and b_j^l are the link biases of the present layer. Nj is the representation of the input feature map. A pooling process is usually performed after a convolution operation. This technique is used to reduce the dimension of the data by down sampling the feature maps. The main concept of the pooling operation is based on the data's local correlation property. Standard pooling is divided into two types that are widely used: maximum pooling and average pooling. Because of its reduced computing complexity, max pooling is more often executed than average pooling, as seen in the following equation.

$$F_i^j = \max_{i \in g_i} (F_i^{j-1}) \tag{8}$$

where max(.) is the argument max function and g_j is the feature map.

After the extraction of the feature, the maps are then fed to a classifier. A classier can be either a DNN or a SVM, any classifier that can handle the data points. A fully connected network is one of the most suitable and widely used classifiers. The target labels are fed to the fully connected layer for loss calculation and back-propagation for training all the layers of the CNN. Finally, a classification operation such as softmax is used to perform class weight normalization and prediction.

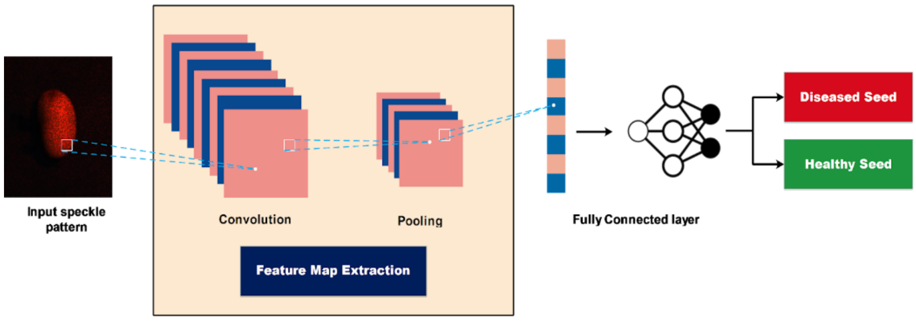


Figure 3 General Architecture of Convolutional Neural Network

4.3 Transfer Learning

Moreover, in the past few years, the pre-trained models have been highly used. The major benefits of TL are fewer data points for retraining the model for the desired objective and improved accuracy. The retraining of the model, training the classifier according to the extracted features, takes much fewer sample data points and less training time (Chen et al., 2020). The application of TL-based algorithms is used in various fields (Zhuang et al., 2019), such as collaborative recommendation (Pan, 2016), activity recognition (Cook et al., 2013), visual categorization (Ling Shao et al., 2015), and others. Several learning strategies have been presented in various industrial applications, including domain adaptive learning (Liu et al., 2020), adversarial learning (Liu et al., 2019), and ensemble deep kernel learning (Liu et al., 2018).

A pre-trained CNN model has been trained on a broad class of datasets in the TL-based technique. The basic block of TL architectures includes an input layer, a convolutional layer followed by a classifier such as DNN, SVM, or a decision tree, as shown in Fig. 4. A CNN trained from scratch requires more computational time, computational power, and a larger dataset. However, in the TL-based approach, the model is a pre-trained model trained on the ImageNet Dataset which has 1000 classes of generic images. By this the kernels are tuned to extract general features, later these are applied to the speckle pattern. ImageNet (n.d.) is a visual database arranged according to the WordNet hierarchy with hundreds of hundreds of pictures for each node. This image dataset is open for all researchers and developers to use. For TL, the models used included; Alexnet (Krizhevsky, 2014), GoogleNet (Szegedy et al., 2015), ResNet18 (He et al., 2016), VGG16 (Simonyan & Zisserman, 2014), and MobileNetV2 (Sandler et al., 2018). Once a pre-trained model is acquired, the weights of the convolutional layers are frozen, i.e. they become non-trainable parameters. Furthermore, the classifier model of the pre-trained model is replaced with a fully connected network or any other classifier which has random weights and can be trained from scratch. The classifier is then trained on the features extracted from the laser backscattering pattern along with the labels provided. As the features extracted are generic in nature, the re-trained of the model takes lesser time and lesser dataset samples. By exploiting this fundamental concept of TL, the re-trained model is able to handle more generalization and to avoid overfitting even by using fewer dataset samples.

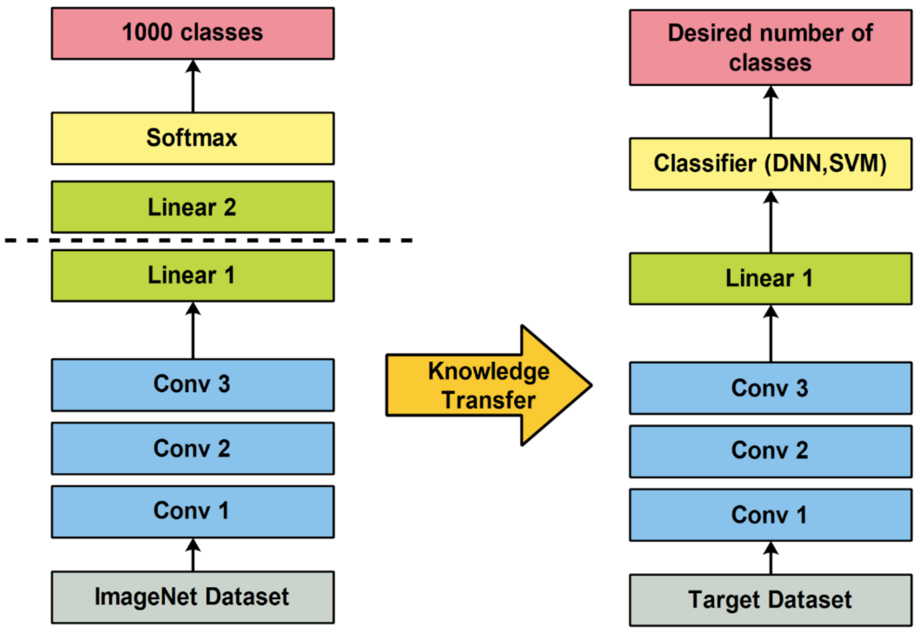


Figure 4 General workflow of transfer learning model

4.4 Ensemble Learning

Most of the techniques, use a single approach, where the proposed processing workflow consists of a single thread. The major concern of such single-thread workflow is that they fail to cover all possible ground and lack the ability to use multiple learners in the case of conventional DL techniques. Hence for many years, an ensemble of learners has gained wide popularity over a single learner. According to (Dietterich, 1997), there is a major three reason for ensemble learning being superior to conventional single learning techniques; firstly, the training data set cannot provide sufficient and enough information for a single selective best learner. Secondly, the search process of selection of learner is arbitrary, an individual cannot be sure if a learner is perfect for a particular task and available dataset. Finally, the target function may not always be present in the hypothesis space being explored. Hence the ensemble-based learning approach is chosen for two major purposes, training the weak learner or feature in the ensemble and having a selection of a strong group of learners or features for better performance. There exist many ensemble strategies such as, averaging method, weighted averaging method, majority vote, winner takes all, early fusion, and many other. These techniques are chosen as per the task and nature of the dataset.

In our case, we have chosen a majority vote and early fusion.

4.4.1 Majority vote-based method

In the method, each of the five TL models was trained on the same dataset. Each of the TL models had a DNN as its classifier. The deep features were extracted from each of the pre-trained models and these features were fed to respective classifiers. After acquiring the final prediction from all the models, the prediction with majority votes, i.e., the statistical model of the tensor containing all the predictions was chosen as the overall prediction of the majority vote-based method. The same is demonstrated in Fig. 5. The models used were Alexnet, GoogleNet, ResNet18, VGG16 and MobileNetV2.

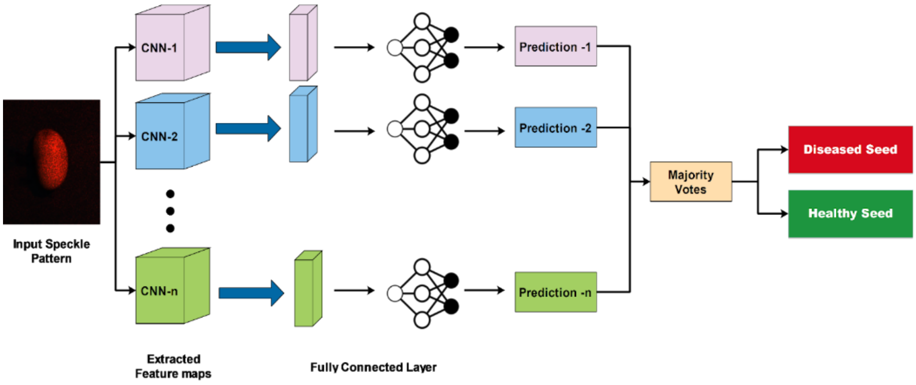


Figure 5 Architecture Overview for majority vote ensemble learning

4.4.2 Early Fusion-based method

In this method, deep features are extracted from the pre-trained models and concatenated in a linear fashion. After concatenation, these deep features are fed to a classifier. A classifier can be either an SVM classifier or a DNN model. The general flow is shown in Fig.6. As there is multiple deep feature feed to the classifier, these deep features cover more ground than individual learners would have covered. Hence the concatenated deep feature provides enough information. This results in increasing the accuracy of the ensemble model as compared to the single learner deep learning model. The criteria for the selection of TL models of early fusion are the individual model's accuracy. Therefore, models having accuracy above 90% were selected in this case, which are, ResNet18, GoogleNet, and MobileNetV2.

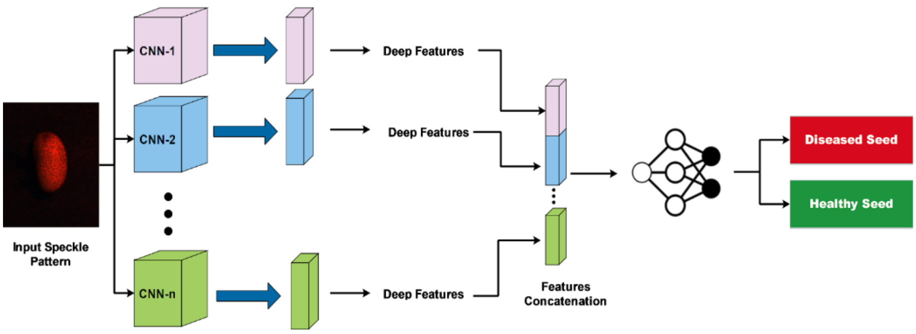


Figure 6 Architecture Overview for early fusion ensemble learning

4.5 Evaluation parameters

We also employed the confusion matrix (Rodriguez et al., 2010) to evaluate the model's performance by using typical assessment criteria such as classification precision, F-1 score, accuracy, specificity, and recall. True positive, true negative, false negative, and false-positive are all entries in a confusion matrix. The evaluation parameters are generated using these four values as shown in the following equations.

$$Accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{9}$$

$$Precision(\%) = \frac{TP}{TP + FP} \times 100 \tag{10}$$

$$Specificity(\%) = \frac{TN}{TN + FP} \times 100 \tag{11}$$

$$Recall(\%) = \frac{TP}{TP + FN} \times 100 \tag{12}$$

$$F - 1\ score(\%) = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{13}$$

Chapter 5

Proposed Deep Learning-based Processing Framework

The suggested DL-based processing and analysis framework will be discussed in this chapter. This chapter will include transfer learning models and an ensemble model, as well as an ensemble learning-based model that is offered as a better alternative to the traditional DL method.

As mentioned earlier, the proposed techniques in the literature survey possess some processing- and experimental-related drawbacks. This motivates us to use a DL-based approach. Moreover, in principle, the use of conventional DL algorithms in our case from scratch trained CNN, lacks the generic nature of the convolutional kernels and requires more time and dataset sample for the training of the model. Hence, the approach using transfer learning and ensemble learning concept using employed.

5.1 Transfer learning models

The models used in this case were obtained from torchvision's package, including models as a sub package (*Models and Pre-Trained Weights—Torchvision Main Documentation*, n.d.). This package contains several pre-trained models. The models have been trained on the ImageNet dataset for 1000 generic classes. The models, being open-source, are widely used by researchers and developers. The average accuracy of these models listed is 50–57% for 1000 class predictions. The convolutional layer's kernel weights are frozen, and the pre-trained classifier is replaced with a DNN or SVM classifier, which will be trained from scratch on the dataset obtained to perform the desired task. Our speckle pattern dataset was used to train a three-layer DNN with [128, 8, 2] nodes, followed by a softmax classifier. The DNN classifier was trained for its respective model's feature maps. The deep CNN-based models are: AlexNet, VGG16, ResNet18, GoogleNet and

MobileNetV2. The overall process workflow for the transfer learning and re-training model on the target dataset is shown in Figure 4. Each of the transfer learning models deployed are briefly described in the subsection below.

5.1.1 AlexNet

The AlexNet Architecture is seen in Figure 7. AlexNet has one of the basic architectures among all the TL models. It includes an input layer followed by a 5-convolutional layer having 55 and 33 kernel size. The extracted features are then fed to a 2 layer fully connected with a softmax classifier at the end. The feature maps are normalized based on the nearest neighbourhood values. AlexNet benign is one of the first and simple CNN models to outperform LeNet (Lecun et al., 1998), which has been widely used since the 1990s.

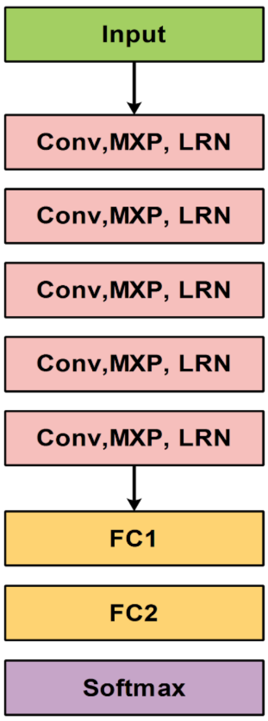


Figure 7 The architecture of AlexNet

5.1.2 VGG16

The Visual Geometry Group (VGG) was proposed in the same year along with AlexNet in 2014. The main feature of VGG was the use of a deep network of critical components to compete against the existing CNNs. The VGG convolutional block employs a convolution operation, ReLU, and a max-pooling operation. The number of blocks utilized varied according to the computing complexity required. The number suffix at the end of the model's name represents the number of convolutional blocks in the convolutional layer. The model employes contains 16 convolution blocks followed by a fully linked network.

Figure 8 The architecture of VGG16

5.1.3 GoogleNet

GoogleNet was the best of the CNN at the time of proposal in 2014. The key feature of GoogleNet was the inclusion of an inception layer with varying filter sizes. A spare data feature map stack is created as a result of this. Furthermore, GoogleNet has a significantly lower number of parameters than AlexNet and VGG19. The initial concept of the Inception Layer is improvised and used GoogleNet The original Inception Layer idea was improved and was employed in GoogleNet for greater performance. The inclusion of an 1×1 convolution operation distinguishes the modified inception model from the naïve inception model.

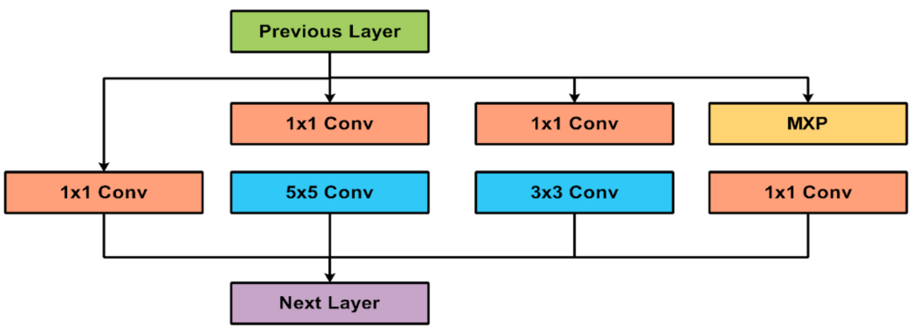


Figure 9 Modified Inception Block used in GoogleNet

5.1.4 ResNet18

ResNet is an abbreviation for Residual Network. ResNet solves the challenge of constructing ultra-deep networks suffering from vanishing gradients. ResNet comes in a variety of versions with 34, 50, 101, 152, and even 1202 layers. Widely is widely utilized in ResNet 50 for higher computational complexity and ResNet 18 for reduced computational complexity with some trade-offs in model performance. Figure 10 shows the basic block diagram. ResNet takes both raw and processed data from the previous layer and sends it on to the next. As a result, the design is referred to as Residual Network.

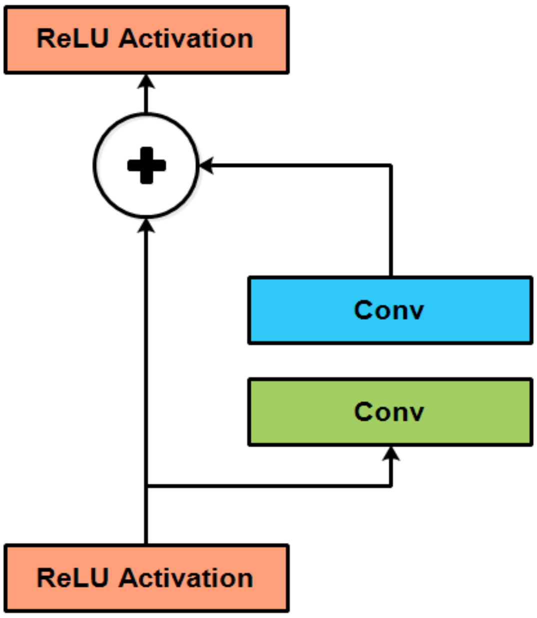


Figure 10 Basic block of ResNet's Residual Block

5.1.5 MobileNetV2

MobileNetV2 is an improved state-of-the-art mobile model that has been demonstrated to operate in many tests and benchmarks. MobielNetV2 was proposed for the purpose of object detection. The biggest disadvantage previously encountered was the non-linear filter, resulting in a lightweight depth wise convolution intermediate expansion layer. The decoupling of input-outputs technique was also utilized to reduce non-linearity in thin layers. This MobileNet model is also available in a variety of sizes and architects, with variable kernel sizes. Fig. 11. Demonstrates the MobileNet basic blocks, which comprise a convolution and ReLU block, depth wise lightweight layer convolution with ReLU, and a convolutional block.

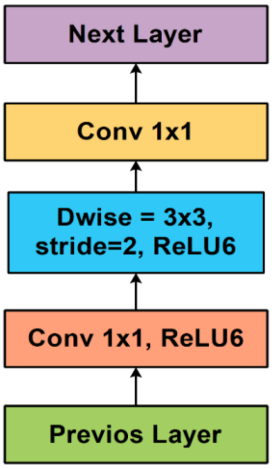


Figure 11 The basic block of MobileNetV2 architecture

Other models, such as InceptionV3, MNASnet, and others, were tested. However, the training accuracy did not increase as predicted due to the model's features and complexity. For ensemble learning, just the five models stated above are employed.

5.2 Ensemble Model and Learning

After the retrained of the model for the targeted dataset, these models are used in ensemble algorithms. Ensemble Model using Majority Vote and Ensemble Learning using early fusion algorithms were used.

5.2.1 Ensemble Model Majority Vote

This algorithm uses the simple concept of voting or polling. The prediction from all the classes is considered for the poll. The models used are completely (re-)trained models having frozen convolutional layer kernels. Finally, the class with the majority vote is considered as the overall prediction of the ensemble model. The architecture of the majority vote ensemble model is shown in Fig. 12. All of the five models, AlexNet, VGG16, ResNet18, GoogleNet and MobileNet are used. The main benefit of this approach is that the algorithm does not require any training. So, the only requirement for this model is to have a trained model for the target dataset and apply the mode function on the prediction output array.

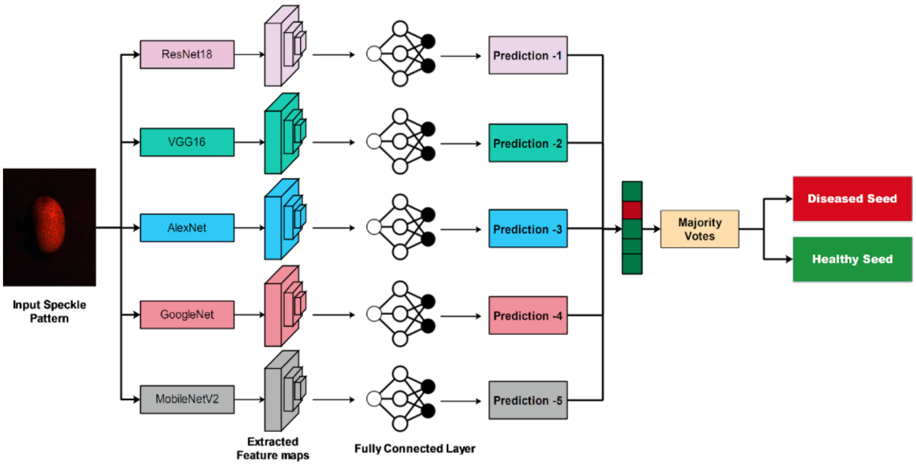


Figure 12 Architecture for Majority Vote Ensemble Model

5.2.2 Ensemble Learning Early Fusion

The majority vote algorithm uses fully (re-)trained models, on the other hand, in the early fusion ensemble learning algorithm only the pre-trained convolution layers are used from different transfer learning models. Among the five transfer learning models, the top three models having the highest accuracy are used; ResNet18, GoogleNet, and MobileNet having testing accuracy of 95.33, 94.67, and 93.67 respectively. The feature extracted from these models is concatenated and forwarded to a fully connected network. A feature is used to train this FC network with targeted labels. The architecture used for ensemble learning is shown in Fig. 13. The key feature of this algorithm is that it provides sufficient data to the fully connected network for training proper and enough data

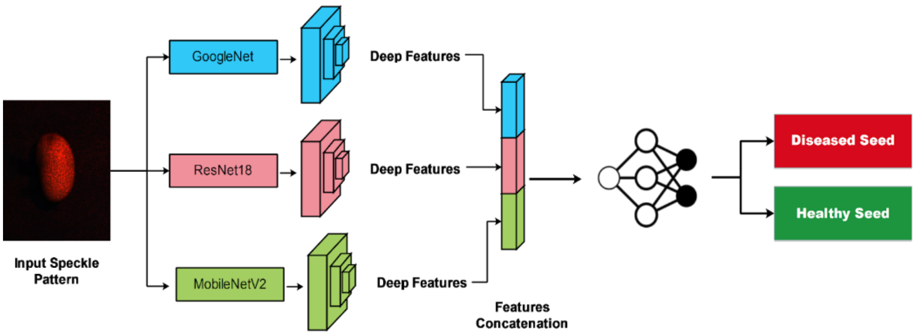


Figure 13 Architecture for Early Fusion Ensemble Learning

5.3 CNN trained from scratch

A CNN model is trained from scratch to be used as the baseline for comparison between different TL-based models, ensemble model, and ensemble learning-based early fusion model. As this model is trained for a specific task, i.e., classification of healthy and diseased seeds using speckle images, the kernels in the convolutional layer are tuned to this specific task only. Kernels in TL-based models, on the other hand, are trained on the ImageNet dataset, which includes 1000 classes, and so are generic in nature. The architecture used for CNN is shown in Fig. 13. The architecture is kept as straightforward as possible due to the simple nature of the dataset. The convolutional layer includes 2 convolutional blocks which include a convolution operation followed by the ReLU activation function and finally max-pool operation. [2,1,3,3] is the kernel size of 1st convolutional layer and [4,1,3,3] is the kernel size for 2nd convolutional layer. For max pooling, 22 kernel size is used. After vector flattening, the data is fed to a fully connected network. The network comprises of 3 layers having [120, 60, 2] nodes respectively. Finally, the softmax function is used to determine the likelihood of the input sample belonging to the target class using the output link weights.

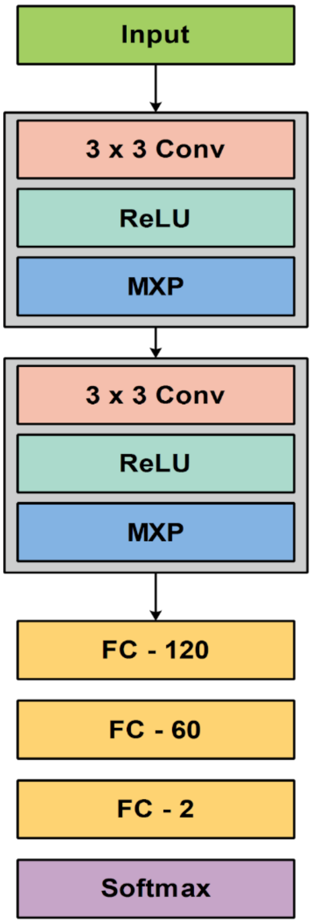


Figure 14 Architecture for CNN trained from scratch

Chapter 6

Results and Discussions

A summary of the target dataset, as well as a brief overview of the work's findings, are included in this chapter. After evaluating numerous models, the best design is suggested based on accuracy and other criteria.

6.1 Speckle Pattern Dataset

The speckle images obtained from the hardware setup in Fig. 2 are divided into two groups: healthy seed samples and damaged seed samples. These speckle patterns are captured are 32fps for 1080 1920 pixels size. After the acquisition of the speckle patterns, the images are cropped to match the size of the seed sample, i.e., 300 300 pixels. After capturing all of the images, labels for the various seed samples are established, such as class: 0 for diseased seeds and class: 1 for healthy samples. Then the images are normalized to a 0 - 255 intensity range for 244×244 pixel size. A total of 1700 images were taken, which were divided into three groups in the ratio of 70:20:10 training, validation, and testing. In the training group, there are 1200 photos with 600 speckle samples for each class, while the validation and testing groups have 300 and 200 images, respectively.

The models used are fine-tuned using the above-mentioned dataset as the target dataset. The models were (re-)trained for 100 epochs using the following parameters, Adam-optimiser, 0.001 as the learning rate, and binary-cross-entropy function for calculation of loss function. Further parameters related to the model are shown in Table 1.

	CNN	ResNet18	VGG16	MobileNetV2	GoogleNet	AlexNet	Ensemble MV	Ensemble EF
Epochs	100	100	100	100	100	100	100	100
Learning Rate	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Batch size	25	25	25	25	25	25	25	25
Loss Function	BCE	BCE	BCE	BCE	BCE	BCE	BCE	BCE
Optimiser	Adam	Adam	Adam	Adam	Adam	Adam	Adam	Adam
Total Parameters	405 K	11 M	139 M	2 M	6 M	61 M	219 M	19 M
Trainable Parameters	405 K	66 K	4 M	165 K	132 K	4 M	0	5 M
Non-Trainable Parameters	0	11 M	134 M	2 M	6 M	57 M	210 M	19 M

Table 1 Hyper parameter details for all the models

6.2 Model Performance Comparison

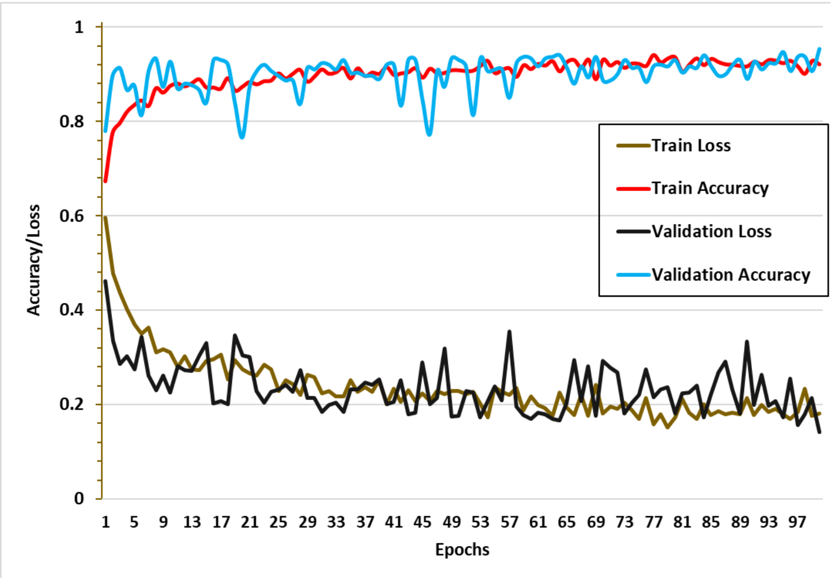
All models are (re-)trained on the same targeted dataset. Various types of TL-based models, ensemble model-based majority vote, ensemble learning-based early fusion model, and lastly, a CNN trained from scratch is used to compare the performance of all the models. Among all the models and algorithms, the ensemble learning-based early fusion model has the highest test accuracy of 98.08%, and VGG16 has the lowest test accuracy of 84.33%. This is presumed to be because the models used for early fusion have the highest individual test accuracy among the five TL-based models. Furthermore, the accuracies of the remaining models are shown in Table 2.

	Accuracy	Precesion	Specificity	Recall	F-1 Score
CNN	95.5	97.48	96.44	98.61	98.05
ResNet18	95.33	96.85	95.77	96.4	96.63
VGG16	84.33	88.45	84.2	94.87	91.55
MobileNetV2	93.67	94.6	93.1	92.78	93.69
GoogleNet	94.67	94.7	92.7	98	96.33
AlexNet	88.9	92.91	92	95.1	94
Ensemble MV	94	92.28	95.2	94.9	93.58
EnSemble EF	98.08	98.98	99	97.17	98.07

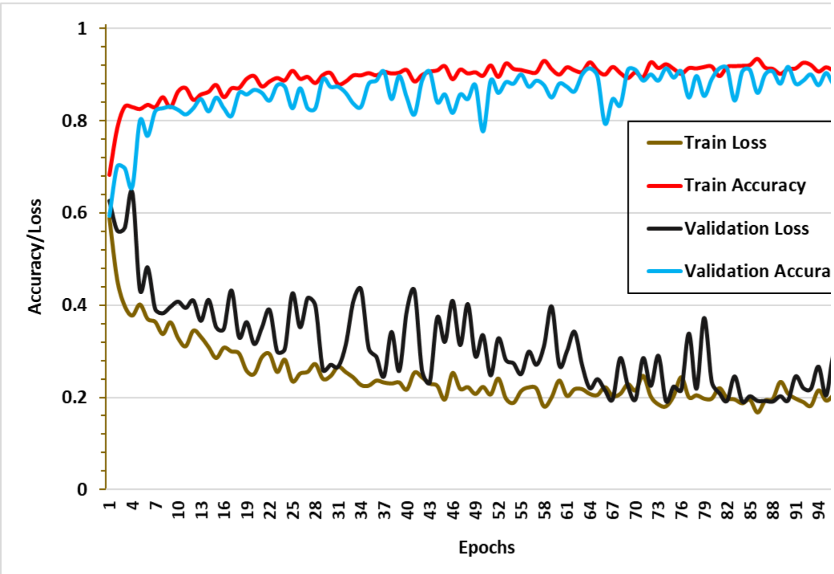
Table 2 Evaluation parameters of all models

Individual model performance is measured by training and validation accuracy, as well as a loss graph. Training and validation accuracies are projected to grow steadily as the number of epochs increases, while training and validation loss is expected to decrease and approach zero. However, as irregularities emerge during model training, there are occasional oscillations in accuracy and loss curves. This might be the result of a divergence from the ideal spot. It is also projected to converge after a few epochs and stay stable as the number of epochs grows.

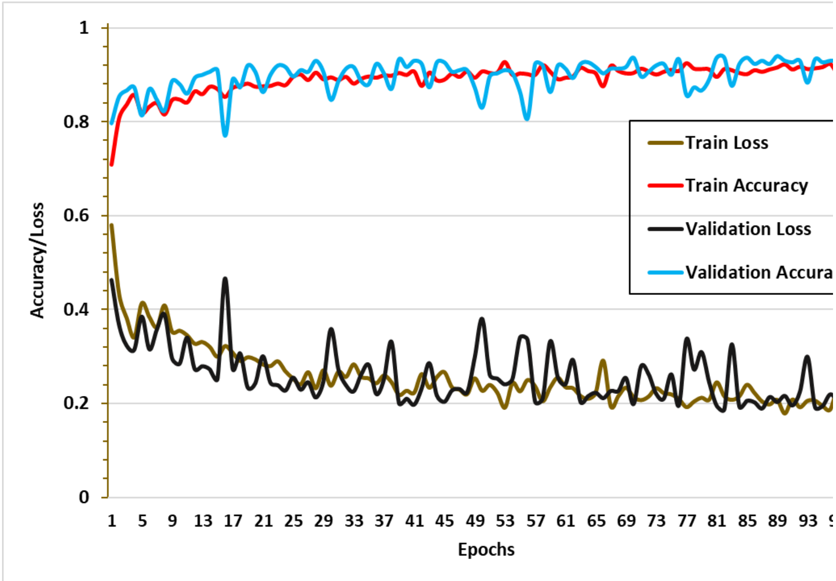
The evaluation parameters and model performance curves for TL-based are shown in Table 2 and Fig. 15. Table 2 contains brief results of all the TL-based model evaluation parameters. These parameters provide an insight into the confusion matrix of the model. Among all the model's total parameters as shown in Table 1, VGG16 has the highest number of total parameters accounting for 13,85,88,122 but provides the least accuracy 84.33% among all the models. On the other hand, the model with the highest accuracy, 95.33% has 1,12,43,226 only. But for a model with the least number of total parameters, GoogleNet, has the second-highest accuracy of 94.67. Hence it can be seen that models with a high number of total parameters provide less accuracy. Moreover, the CNN trained from scratch has an accuracy of 95.5 which is more than all of the TL-based models with total parameters of 4,05,038 which accounts for only 3% of the total number of parameters used in ResNet18. The performance plots for CNN trained from scratch can be seen in Fig. 16.



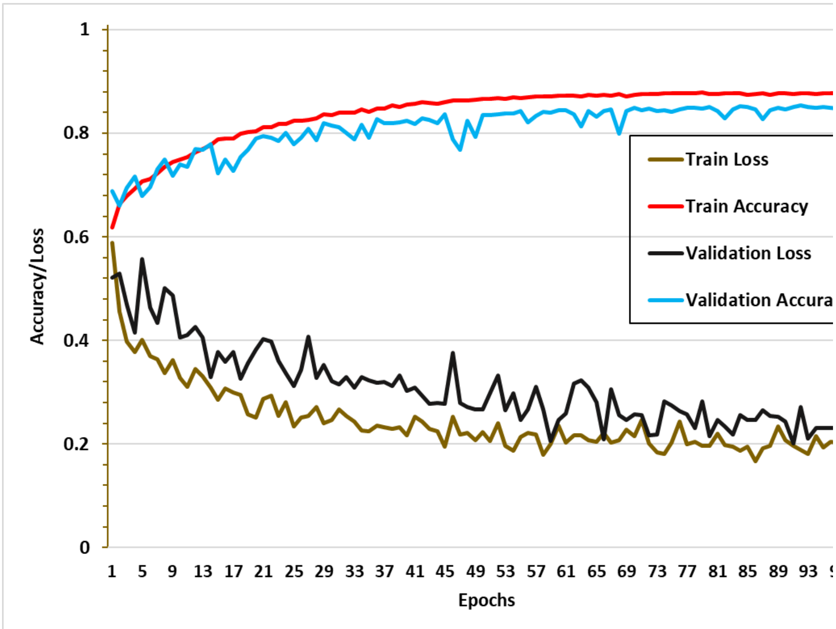
(a)



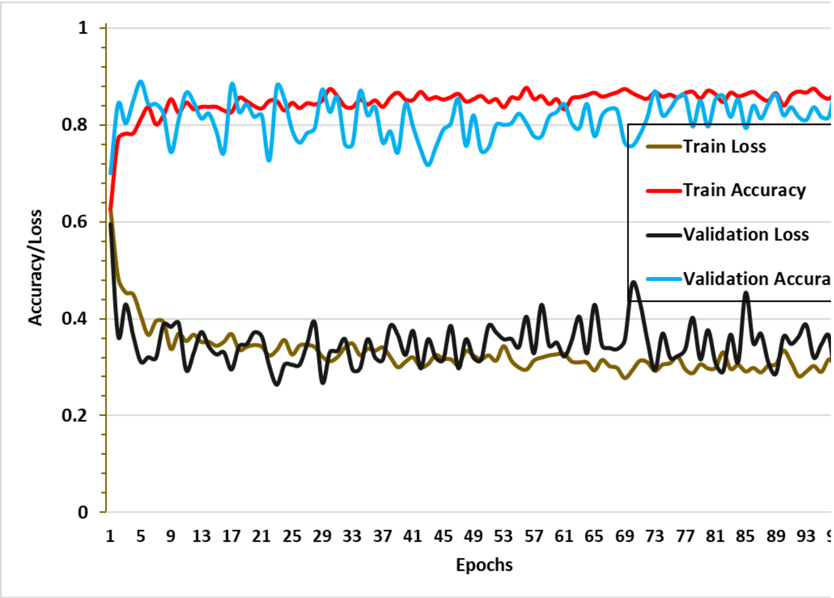
(b)


1
0.8
0.6
0.4
0.2
0
Accuracy/Loss
Train Loss
Train Accuracy
Validation Loss
Validation Accura
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93
Epochs

(c)


1
0.8
0.6
0.4
0.2
0
Accuracy/Loss
Train Loss
Train Accuracy
Validation Loss
Validation Accura
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93
Epochs

(d)



(e)

Figure 15 Model Performance curve for (a) ResNet18, (b) MobileNet, (c) GoogleNet, (d) VGG16 and (e) AlexNet18

After achieving the accuracy for TL-based models and CNN model, the Ensemble model, and Ensemble learning approach was implemented. The ensemble model uses a majority vote algorithm. This algorithm uses all of the five models' predictions and the class majority votes will declare the final prediction class. ResNet18, GoogleNet, MobileNetV2, AlexNet, and VGG16 are used which gives an accuracy of 96.1% which is more than all the individual models. Above all, because retrained models are employed, no training is necessary. Hence the trainable parameter in this algorithm is zero. The downside of this algorithm is the requirement of being completely trained on the targeted dataset of a similar dataset (Turkoglu et al.,2021).

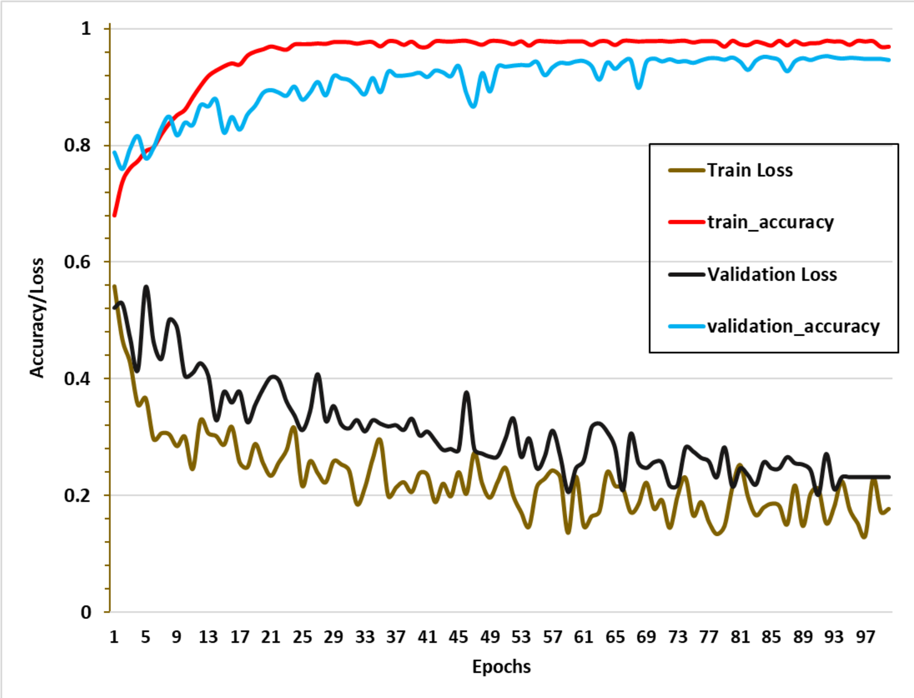


Figure 16 Model Performance curve for CNN

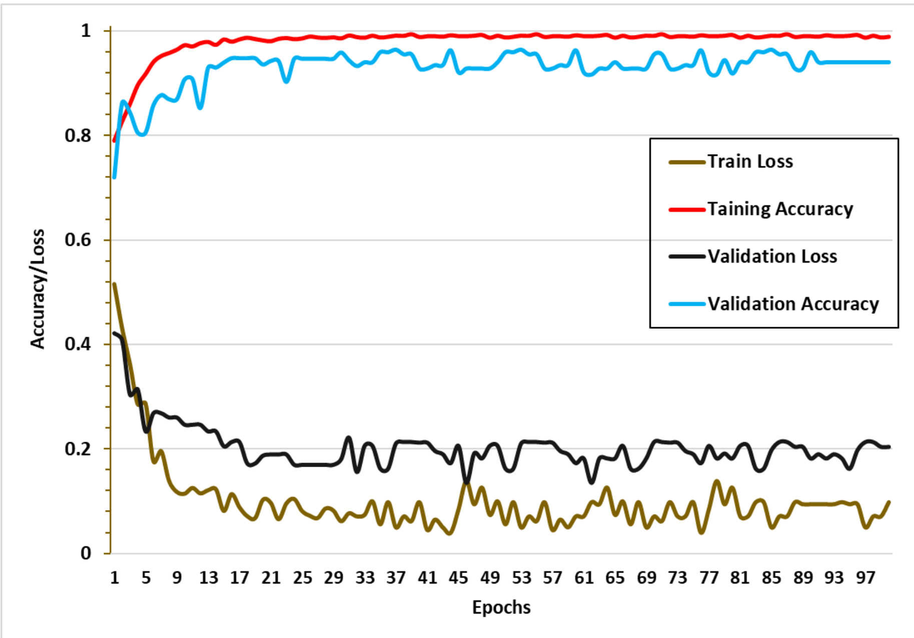


Figure 17 Model Performance curve for EF-based model

Finally, there is the ensemble learning-based approach, which requires training. The models' only convolutional layer is taken and the features given by each of the layers are concatenated. Because only convolutional layers with frozen weights are employed, the number of non-trainable parameters is increased. Convolutional layers from ResNet18, GoogleNet, and MobileNetV2 models are utilized because they have the best accuracy in individual model training. Fig. 16 shows the model's performance curve for the early fusion-based model. With a score of 98.08, this model's accuracy is the greatest of all the proposed models and algorithms in this study. Although this technique has a large number of parameters, the total number of trainable parameters is just 45,89,746.

Chapter 7

Conclusion and Future Work

The proposed work's conclusion, as well as possible future scope for expanding this work to additional applications and domains.

In this thesis, a deep learning-based optical techniques analytical pipeline using speckle patterns to analyze and identify diseased and healthy seed samples is presented. The use of deep learning-based techniques has proven to be more advantageous than traditional techniques, which rely on manual operations such as ROI marking, feature extraction such as edges or visual descriptors, and statistically calculated values that are dependent on several image quality factors such as contrast, roughness, or image background activity, among others. Furthermore, the usage of TL improves the necessity for a smaller dataset for model (re-)training. Furthermore, the kernel of the convolutional layers of TL models has a general character that yields superior features in theory. When compared to a CNN trained from scratch for a specific task, the CNN surpasses the TL-models by 1-2% accuracy. Also, the use of ensemble-based models uses the advantage of both DNN and TL. The key feature of ensemble learning algorithms is that the strong learner overcomes the drawback of the weak learner. Therefore, as proven by the acquired results, both the ensemble-based models have higher test accuracy than the CNN and TL-based models. As for highlights of this work, it can be said that the process requires a lesser dataset, no maul operations, lesser image quality correction, and high accuracy in the detection of diseased and healthy seed samples due to the use of state-of-the-art techniques.

Future Work

The method proposed is based on TL and ensemble learning, and it employs five models and two algorithms with the use of spatial speckle patterns. The future work, may include further exploitation and exploration of the acquired dataset or acquisition of more speckle image of different biologically active objects to extend the application to include other seeds and crops, as well as different diseases Furthermore, other DL-based models may be researched with regard to the intended goal, such as RNN, and alternative approaches, such as unsupervised learning, can be explored. Also, to reduce the execution time so that the method can be deployed online in the industry.

Reference

- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A. S., & Asari, V. K. (2019). A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics*, *8*(3), 292. https://doi.org/10.3390/electronics8030292
- Alves, J. A., Braga, R. A., Júnior, & Vilas Boas, E. V. de B. (2013). Identification of respiration rate and water activity change in fresh-cut carrots using biospeckle laser and frequency approach. *Postharvest Biology and Technology*, *86*, 381–386. https://doi.org/10.1016/j.postharvbio.2013.07.030
- Arizaga, R. (2002). Display of local activity using dynamical speckle patterns. *Optical Engineering*, *41*(2), 287. https://doi.org/10.1117/1.1428739
- Arizaga, R., Trivi, M., & Rabal, H. (1999). Speckle time evolution characterization by the co-occurrence matrix analysis. *Optics & Laser Technology*, *31*(2), 163–169. https://doi.org/10.1016/s0030-3992(99)00033-x
- Braga, R. A., Cardoso, R. R., Bezerra, P. S., Wouters, F., Sampaio, G. R., & Varaschin, M. S. (2012). Biospeckle numerical values over spectral image maps of activity. *Optics Communications*, *285*(5), 553–561. https://doi.org/10.1016/j.optcom.2011.10.079
- Braga, R. A., Nobre, C. M. B., Costa, A. G., Sáfadi, T., & da Costa, F. M. (2011). Evaluation of activity through dynamic laser speckle using the absolute value of the differences. *Optics Communications*, *284*(2), 646–650. https://doi.org/10.1016/j.optcom.2010.09.064
- Briers, J. D. (1996). Laser speckle contrast analysis (LASCA): a nonscanning, full-field technique for

monitoring capillary blood flow. *Journal of Biomedical Optics*, *1*(2), 174. https://doi.org/10.1117/12.231359

- Cardoso, R. R., Costa, A. G., Nobre, C. M. B., & Braga, R. A., Jr. (2011). Frequency signature of water activity by biospeckle laser. *Optics Communications*, *284*(8), 2131–2136. https://doi.org/10.1016/j.optcom.2011.01.003
- Chatterjee, A., Saha, D., Niemann, H., Gryshkov, O., Glasmacher, B., & Hofmann, N. (2017). Effects of cryopreservation on the epigenetic profile of cells. *Cryobiology*, *74*, 1–7. https://doi.org/10.1016/j.cryobiol.2016.12.002
- Chatterjee, A., Singh, P., Bhatia, V., & Prakash, S. (2018). A Low-Cost Optical Sensor for Secured Antispoof Touchless Palm Print Biometry. *IEEE Sensors Letters*, *2*(2), 1–4. https://doi.org/10.1109/lsens.2018.2837879
- Chatterjee, A., Singh, P., Bhatia, V., & Prakash, S. (2019). Ear biometrics recognition using laser biospeckled fringe projection profilometry. *Optics & Laser Technology*, *112*, 368–378. https://doi.org/10.1016/j.optlastec.2018.11.043
- Chatterjee, A., Singh, P., Bhatia, V., & Prakash, S. (2020). An efficient automated biospeckle indexing strategy using morphological and geo-statistical descriptors. *Optics and Lasers in Engineering*, *134*, 106217. https://doi.org/10.1016/j.optlaseng.2020.106217
- Chen, J., Chen, J., Zhang, D., Sun, Y., & Nanehkaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. *Computers and Electronics in Agriculture*, *173*, 105393. https://doi.org/10.1016/j.compag.2020.105393
- Cook, D., Feuz, K. D., & Krishnan, N. C. (2013). Transfer learning for activity recognition: a survey. *Knowledge and*

Information Systems, *36*(3), 537–556. https://doi.org/10.1007/s10115-013-0665-3

- Dhanotia, J., Bopche, L., Bhatia, V., & Prakash, S. (2017). Quality Assessment of Medicinal Leaves Through Biospeckle Technique. In *Springer Proceedings in Physics* (pp. 389–394). Springer Singapore. http://dx.doi.org/10.1007/978-981-10-3908-9_47
- Dietterich, & T.G. (1997). Machine learning research: Four current directions. *AI Magazine*, *18*(4), 97–136.
- Draijer, M., Hondebrink, E., van Leeuwen, T., & Steenbergen, W. (2008). Review of laser speckle contrast techniques for visualizing tissue perfusion. *Lasers in Medical Science*, *24*(4), 639–651. https://doi.org/10.1007/s10103-008-0626-3
- Fujii, H., Nohira, K., Shintomi, Y., Asakura, T., & Ohura, T. (1985). Blood flow observed by time-varying laser speckle. *Optics Letters*, *10*(3), 104. https://doi.org/10.1364/ol.10.000104
- Fujii, H., Nohira, K., Yamamoto, Y., Ikawa, H., & Ohura, T. (1987). Evaluation of blood flow by laser speckle image sensing Part 1. *Applied Optics*, *26*(24), 5321. https://doi.org/10.1364/ao.26.005321
- Godinho, R. P., Silva, M. M., Nozela, J. R., & Braga, R. A. (2012). Online biospeckle assessment without loss of definition and resolution by motion history image. *Optics and Lasers in Engineering*, *50*(3), 366–372. https://doi.org/10.1016/j.optlaseng.2011.10.023
- He, K., Zhang, X., Ren, S., & Sun, J. (2016, June). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. http://dx.doi.org/10.1109/cvpr.2016.90

- Ibrahim, E. A. (2016). Seed priming to alleviate salinity stress in germinating seeds. *Journal of Plant Physiology*, *192*, 38–46. https://doi.org/10.1016/j.jplph.2015.12.011
- *ImageNet*. (n.d.). Retrieved May 22, 2022, from https://www.image-net.org/
- Jayanthy, A. K., Sujatha, N., & Ramasubba Reddy, M. (2011). Non-invasive assessment of static scatterer concentration in phantom body fluids using laser speckle contrast analysis technique. *Optics and Lasers in Engineering*, *49*(4), 553–556. https://doi.org/10.1016/j.optlaseng.2010.12.003
- Kalyzhner, Z., Levitas, O., Kalichman, F., Jacobson, R., & Zalevsky, Z. (2019). Photonic human identification based on deep learning of back scattered laser speckle patterns. *Optics Express*, *27*(24), 36002. https://doi.org/10.1364/oe.27.036002
- Krizhevsky, A. (2014, April 23). *One weird trick for parallelizing convolutional neural networks*. ArXiv.Org. https://arxiv.org/abs/1404.5997
- Kulkarni, R., Pal, P., & Banoth, E. (2021). Spatio-temporal analysis of dynamic speckle patterns using singular value decomposition. *Optics and Lasers in Engineering*, *142*, 106588. https://doi.org/10.1016/j.optlaseng.2021.106588
- Kumari, S., & Nirala, A. K. (2019). Biospeckle image processing algorithms for non-destructive differentiation between maturity and ripe stages of Indian climacteric fruits and evaluation of their ripening period. *Laser Physics*, *29*(7), 075601. https://doi.org/10.1088/1555-6611/ab0c93
- Learning transfer and cognitive development. (1965). *Comprehensive Psychiatry*, *6*(6), 431–432. https://doi.org/10.1016/s0010-440x(65)80065-7
- Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition.

Proceedings of the IEEE, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

- Ling Shao, Fan Zhu, & Xuelong Li. (2015). Transfer Learning for Visual Categorization: A Survey. *IEEE Transactions on Neural Networks and Learning Systems*, *26*(5), 1019–1034. https://doi.org/10.1109/tnnls.2014.2330900
- Liu, Y., Yang, C., Gao, Z., & Yao, Y. (2018). Ensemble deep kernel learning with application to quality prediction in industrial polymerization processes. *Chemometrics and Intelligent Laboratory Systems*, *174*, 15–21. https://doi.org/10.1016/j.chemolab.2018.01.008
- Liu, Y., Yang, C., Liu, K., Chen, B., & Yao, Y. (2019). Domain adaptation transfer learning soft sensor for product quality prediction. *Chemometrics and Intelligent Laboratory Systems*, *192*, 103813. https://doi.org/10.1016/j.chemolab.2019.103813
- Liu, Y., Yang, C., Zhang, M., Dai, Y., & Yao, Y. (2020). Development of Adversarial Transfer Learning Soft Sensor for Multigrade Processes. *Industrial & Engineering Chemistry Research*, *59*(37), 16330–16345. https://doi.org/10.1021/acs.iecr.0c02398
- Ministry of Statistics and Programme Implementation. (2021, May 31). *Provisional estimates of Annual National Income, 2020-21 and Quarterly estimates (Q4) of Gross Domestic Product, 2020-21*. https://pib.gov.in/PressReleasePage.aspx?PRID=1723153
- *Models and pre-trained weights — Torchvision main documentation*. (n.d.). Retrieved May 24, 2022, from https://pytorch.org/vision/main/models.html
- Oulamara, A., Tribillon, G., & Duvernoy, J. (1989). Biological Activity Measurement on Botanical Specimen Surfaces Using a Temporal Decorrelation Effect of Laser

Speckle. *Journal of Modern Optics*, *36*(2), 165–179. https://doi.org/10.1080/09500348914550221

- Pajuelo, M., Baldwin, G., Rabal, H., Cap, N., Arizaga, R., & Trivi, M. (2003). Bio-speckle assessment of bruising in fruits. *Optics and Lasers in Engineering*, *40*(1–2), 13–24. https://doi.org/10.1016/s0143-8166(02)00063-5
- Pan, W. (2016). A survey of transfer learning for collaborative recommendation with auxiliary data. *Neurocomputing*, *177*, 447–453. https://doi.org/10.1016/j.neucom.2015.11.059
- Rajput, L. S., Madhusudhan, P., & Sinha, P. (2020). Epidemiology of seed-borne diseases. In *Seed-Borne Diseases of Agricultural Crops: Detection, Diagnosis & Management* (pp. 863–871). Springer Singapore. http://dx.doi.org/10.1007/978-981-32-9046-4_30
- Rodriguez, J. D., Perez, A., & Lozano, J. A. (2010). Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 569–575. https://doi.org/10.1109/tpami.2009.187
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018, June). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. http://dx.doi.org/10.1109/cvpr.2018.00474
- Simonyan, K., & Zisserman, A. (2014, September 4). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. ArXiv.Org. https://arxiv.org/abs/1409.1556
- Singh, P., Chatterjee, A., Bhatia, V., & Prakash, S. (2020). Application of laser biospeckle analysis for assessment of seed priming treatments. *Computers and Electronics in Agriculture*, *169*, 105212. https://doi.org/10.1016/j.compag.2020.105212

- Sood, S., Mahajan, S., Doegar, A., & Das, A. (2016, September). Internal crack detection in kidney bean seeds using X-ray imaging technique. *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. http://dx.doi.org/10.1109/icacci.2016.7732388
- Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015, June). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. http://dx.doi.org/10.1109/cvpr.2015.7298594
- Turkoglu, M., Yanikoğlu, B., & Hanbay, D. (2021). PlantDiseaseNet: convolutional neural network ensemble for plant disease and pest detection. *Signal, Image and Video Processing*, *16*(2), 301–309. https://doi.org/10.1007/s11760-021-01909-2
- Xuan, Q., Fang, B., Liu, Y., Wang, J., Zhang, J., Zheng, Y., & Bao, G. (2018). Automatic Pearl Classification Machine Based on a Multistream Convolutional Neural Network. *IEEE Transactions on Industrial Electronics*, *65*(8), 6538–6547. https://doi.org/10.1109/tie.2017.2784394
- Yang, J., Wang, J., Lu, G., Fei, S., Yan, T., Zhang, C., Lu, X., Yu, Z., Li, W., & Tang, X. (2021). TeaNet: Deep learning on Near-Infrared Spectroscopy (NIR) data for the assurance of tea quality. *Computers and Electronics in Agriculture*, *190*, 106431. https://doi.org/10.1016/j.compag.2021.106431
- Yao, J.-P., Wang, Z.-Y., de Oliveira, R. F., Wang, Z.-Y., & Huang, L. (2021). A deep learning method for the long-term prediction of plant electrical signals under salt stress to identify salt tolerance. *Computers and Electronics in Agriculture*, *190*, 106435. https://doi.org/10.1016/j.compag.2021.106435

- Zdunek, A., Adamiak, A., Pieczywek, P. M., & Kurenda, A. (2014). The biospeckle method for the investigation of agricultural crops: A review. *Optics and Lasers in Engineering*, *52*, 276–285. https://doi.org/10.1016/j.optlaseng.2013.06.017
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2019, November 7). *A Comprehensive Survey on Transfer Learning*. ArXiv.Org. https://arxiv.org/abs/1911.02685