Deep Learning for Image Time-Series Analysis: Application to Crop Mapping

MEGHA SOOD
September, 2021

SUPERVISORS:
Dr. Anil Kumar, IIRS
Dr. Claudio Persello, ITC

Deep Learning for Image Time-Series Analysis: Application to Crop Mapping

MEGHA SOOD
Enschede, The Netherlands, September, 2021

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:
Dr. Anil, Kumar, IIRS
Dr. Claudio, Persello, ITC

THESIS ASSESSMENT BOARD:
Dr. Alfred Stein, ITC (Chair)
Dr. R.D. Garg (External Examiner, IIT, Roorkee)

ABSTRACT

In the present times, with the dynamically changing climatic conditions, it becomes very important to monitor and estimate the acreage and production of crops to meet the requirements of the bulging population. Conventionally, crop classification was carried out using unsupervised and supervised methods which demanded a lot of manpower and time investment. This research explores the automation capability of deep learning models for finding hidden patterns in the dataset. Keeping this in mind, an optimized number of dates were selected for generating training data for the deep learning model. 1-Dimensional Convolutional Neural Networks (CNN) and its integration with Long Short Term Memory (CNN-LSTM) models were explored and optimized to produce classified outputs. The Modified Possibilistic c-Means (MPCM) Classifier was utilized to test its performance in terms of handling intra-class variability within a crop class and compared with the deep learning model outputs. Also, the dual-sensor approach was studied to find any additional information useful for classification.

The study area chosen was Ludhiana district in the agricultural state of Punjab, India. The research was divided into two phases with maize (Zea Mays), mentha (Mint), and guava (Psidium guajava) as target crops in phase-1 while early potato (Solanum Tuberosum), mid potato, and late potato as target crops in phase-2. Time series data was used to resolve the issue of overlapping spectral signatures of crops. The freely available Sentinel-2 dataset integrated with the Landsat-8 dataset was used in phase-1 of the study while the Sentinel-2 dataset was utilized in the phase-2 study. Class-Based Sensor Independent Normalized Differential Vegetation Index (CBSI-NDVI) was used to enhance the target crop and separate it from the other non-interest crops in the study area. The optimum number of dates was selected for each target crop by performing separability analysis. The date combinations that gave maximum separation between the target crop class and other (non-interest) crop classes were selected to be optimum. Then the dataset generated was used to create training data for the deep learning models. Two deep learning models were explored and optimized which were the 1D-CNN model and an integrated CNN-LSTM model. The deep learning models handle the heterogeneity within a crop class (intra-class variability) very well which is still a challenge for a fuzzy classifier. Here, a fuzzy MPCM classifier was studied against the training data to test its capability of handling the heterogeneity within a crop class during classification and was compared with the outputs of deep learning models. Also, a dual-sensor approach was explored where the Sentinel-1 Synthetic Aperture Radar (SAR) dataset was used alongside optical data to test for any additional information extracted from the data.

It was observed that the CBSI-NDVI index proved to be reliable in enhancing a specific crop in the study area. The optimum dates selected using separability analysis helped to reduce the temporal domain. In phase-1, six-date combinations came out to be optimum for Maize and Mentha while seven-date combinations came out to be optimum for Guava. In phase-2, four-date combinations came out to be optimum for early and mid-potato while three-date combinations came out to be optimum for late potato. A simpler CNN model architecture proved to be efficient for specific crop classification which consisted of two 1D-CNN layers followed by a max pool layer, a dropout layer, and a dense layer. The integrated CNN-LSTM model consisted of three pairs of 1D-CNN and LSTM layers, three dropout layers, two max pool layers, and a dense layer. It was observed that the hybrid model resulted in better classification outputs for the target crops from both phases. It was also observed that the hybrid CNN-LSTM model handled the intra-class variability (heterogeneity within a crop class) best when compared to 1D-CNN and MPCM classifiers. However, the MPCM classifier proved to handle heterogeneity within a crop class better than the CNN model. In the end, it was also observed that using a single sensor dataset yielded better classification results when compared to the dual-sensor dataset.

After analyzing the results, the hybrid CNN-LSTM model provided the best classification results. Also, the fuzzy MPCM classifier performed similarly to a deep learning model in terms of handling heterogeneity within a crop class.

Keywords: *Deep Learning models, Time Series, CBSI-NDVI, Fuzzy based classifier, CNN, CNN-LSTM, MPCM, Crop classification.*

ACKNOWLEDGEMENTS

The entire research period was full of challenges and demanded a lot of support and motivation especially when the whole world was going through a pandemic. Therefore I would like to use this opportunity to thank everyone who has supported and helped me throughout my master's program.

I would like to start by expressing my gratitude to the almighty for presenting me with this opportunity and giving me the courage to pass through all the challenging times.

I would like to express my deepest gratitude to my IIRS supervisor, Dr. Anil Kumar for his constant guidance and support throughout the research period. I would also like to appreciate his motivational words which helped me to keep working and overcome every challenge encountered.

I would also like to express my deepest gratitude to my ITC supervisor, Dr. Claudio Persello for his critical reviews and suggestions throughout the research period. I learned a lot from him which would prove to be beneficial for me in the future.

Also, I am most thankful to our program coordinators, Dr. Sameer Saran and Dr. Sander Oude Elberink who has very patiently coordinated and found solutions to every problem faced by the students during the pandemic. They made sure that no student suffered in any way due to the pandemic by providing help whenever possible. I am also indebted to Dr. Prakash Chauhan, Director, IIRS, Dehradun for his guidance and encouragement throughout the course of my study.

I would also like to thank Dr. Alfred Stein for his constructive criticism of my research.

I would also like to thank my parents and my brother for their constant support and encouragement that helped me get through the challenging times. I would also like to thank my friends and my classmates who always managed to help me lighten up during stressful periods. My family, friends and classmates have supported me both emotionally and mentally, for which I will always be in their debt.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
1D	One Dimensional
2D	Two Dimensional
3D	Three Dimensional
AI	Artificial Intelligence
CBSI	Class-Based Sensor Independent
CNN	Convolutional Neural Networks
DL	Deep Learning
EO	Earth Observation
ESA	European Space Agency
FCC	False Colour Composite
FCM	Fuzzy c-Means
GIS	Geographic Information System
GRD	Ground Range Detected
JM	Jeffries-Matusita
LSTM	Long-short term memory
LULC	Land Use Land Cover
ML	Machine Learning
MLP	Multi-Layer Perceptron
MMD	Mean Membership Difference
MPCM	Modified Possibilistic c-Means
MSI	Multi-Spectral Instrument
NASA	National Aeronautics and Space Administration
NDVI	Normalized Differential Vegetation Index
NIR	Near Infra-Red
NN	Neural Network
OLI	Operational Land Imager
PCM	Possibilistic c-Means
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Networks
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SLC	Single Look Complex
SNAP	Sentinel Application Platform
SWIR	Short Wave Infra-Red
TD	Transformed Divergence
USGS	United States Geological Survey
UTM	Universal Transverse Mercator

1. INTRODUCTION

1.1. Research Background

'Zero Hunger' is the second sustainable development goal that aims at achieving food security around the world (Assembly, 2015). Globally, people have come together and are continuously working in the direction of accomplishing these goals for a better future. There are broadly four targets and indicators for this goal: sufficient availability of food, accessibility to people, adequate consumption of food by people, and stability of the previous three factors to ensure food security. Amongst these factors, the availability of food depends upon the production of crops in a country and acts as an important factor to be monitored, especially in a developing country like India. The government needs to make sure that the crops being produced are healthy for fulfilling the nutritious requirement with good yield.

Earth Observation (EO) is the collection of information about the physical, chemical, and biological systems on the Earth's surface. Remote sensing images have been extensively used for land use/land cover mapping, at level one classification. In remote sensing, Level 1 classification means classifying the image into broader classes like agricultural land, water bodies, urban land, barren land, etc. while in Level 2 classification, classes are sub-divided to form detailed classes, as shown in table 1 (Saritha & Santhosh Kumar, 2015). Level one classification has been implemented using conventional classification techniques such as Support Vector Machines, Random forest, k-means, etc. These classifiers have been used as hard classification with supervised and unsupervised classification modes. Some of the classifiers have been applied as soft classifiers using the fuzzy logic concept (Murmu & Biswas, 2015). The approach adopted in this research is Level-2 and 3 classifications in which different types of crop classes, and, crop sub-classes are mapped using temporal information as crop phonology, respectively.

Table 1: Classification Levels (Saritha & Santhosh Kumar, 2015)

Level 1	Level 2
Builtup Land	Residential
	Commercial
	Industrial
	Transportation
	Mixed Urban land
Agricultural Land	Cropland
	Horticultural
Forest Land	Deciduous Forest Land
	Evergreen Forest Land
	Mixed Forest Land
Barren Land	Sandy Areas Bare
	Exposed Rock
	Play Grounds

For sustainable natural resources, there is a need to manage agriculture by using effective cropland mapping and monitoring (Matton et al., 2015). Over time, there has been a tremendous amount of improvement in the spatial and temporal resolution of the remote sensing datasets. This helps to make use of the remote sensing datasets in the applications such as crop or forest tree species identification, land cover mapping, urban planning and disaster mitigation planning. Remote sensing image classification remains a challenging task when specific agricultural crop classes are considered as there can be spectral overlap between the target class and other classes (Mazzia et al., 2020). In the case of crops, the growth stage of the crops changes with time having a unique phonological profile. This profile can be introduced using multi-temporal images. While using multi-temporal remote sensing data for the crop type mapping, the classifier should be capable of mapping specific classes of interest. The temporal data incorporates the phonological information of the crop to overcome the spectral overlap between the crop classes in the study area. While maintaining spatial (x, y) components of data, two of the three dimensions available are fixed, and there is one dimension left while two components, i.e., spectral or temporal that can be considered. So, the spectral dimension of each image has to be reduced to maintain the temporal dimension. This goal can be achieved by evaluating class-based sensor independent – normalized differential vegetation index (CBSI-NDVI) which has proved to be efficient in crop mapping as well as in identifying fields with damaged crops (Rawat et al., 2020). The second advantage of using the CBSI-NDVI approach is that based on data sets as well as a class type, this approach selects appropriate bands from each temporal image (Misra et al., 2012). Meaning, it chooses the spectral bands that gives the maximum separation between the target crop and non-interest crop classes. Further to include unique changes, a potential solution to the spectral overlapping problem can be by increasing the temporal resolution of the dataset used. While using the multi-spectral datasets like Sentinel-2, it can be taken care to incorporate data from sentinel 2A and 2B series, with 5 days' time intervals. As per the references, it is clear that a large number of fuzzy-based and other classifiers have been used for single-class mapping (Musande, Kumar, & Kale, 2012). But learning-based algorithms have not been tested or applied on temporal data processing for specific class mapping.

The non-parametric classifier like Neural Networks (NN) has proved to perform better when compared to the parametric classifier in those cases where the distribution of data in the area is unknown (Gómez, White, & Wulder, 2016). NN-based algorithms are designed to mimic the human brain. Neural networks with their modified versions such as Convolutional Neural Network (CNN) and Recurrent Neural Network - Long Short Term Memory (RNN-LSTM) are comparatively newer techniques that are being actively explored for the classification purpose from past six years. The layers in a CNN architecture include a convolution, a non-linear activation function, and max pooling (Paola & Schowengerdt, 1995). Initially, when the CNN model is trained, a set of random weights are assigned to the layer which is optimized over the successive layers to minimize the model loss and maximize the model performance. An advantage of using RNN is that it stores the previously predicted values in its memory which is then considered along with the new input to predict the succeeding value. LSTM is a modified version of RNN where it stores the most relevant values in its memory while predicting the future values. Since it considers the values which have more importance over the others and discards other values, it is a faster and better approach (Sun et al., 2019). 1D-CNN has proved to perform better when the training data is less or when an application is developed for a specific objective (Kiranyaz et al., 2019). CNN with its other versions has not been explored much for temporal data processing. Studies have shown that the results from this algorithm are similar to the conventional classifiers such as maximum-likelihood but it still needs further exploration before it could be adopted as an alternative to the conventional classifiers (Paola & Schowengerdt, 1995).

More than a decade ago, fuzzy classifiers were used for specific crop mapping where the output generated was better than the conventional classifiers (Assimakopoulos et al., 2003). The deep learning models have proved to handle the heterogeneity within a class however it is still a challenge for fuzzy classifiers. Therefore, a modified version of the Possibilistic c-Means (MPCM) classifier was tested against the training dataset and the output was compared with the results obtained from the deep learning model. Here, a modified version of the fuzzy c-Means classifier was utilized which was made to perform similar to a deep learning model by considering an individual training data as input without performing statistical evaluation like mean on the inputs data. The output generated from the MPCM algorithm was further explored to handle heterogeneity within a class and compared with the outputs from deep learning models. Also, microwave Synthetic Aperture Radar (SAR) image was incorporated with the optical image to test for any additional information extracted by using dual-sensor datasets for specific crop classification.

1.2. Problem Statement

In India, every year the state government estimates the yield of each crop at the district level and state level. For yield estimation, it is necessary for the government bodies to know the area under cultivation for each crop. Instead of using the traditional methods where each crop information was calculated manually, remote sensing is used for calculating the area under cultivation of each crop, which saves a lot of time and manpower. Several factors contribute to the selection of sowing date for a crop by the farmers like the variety of crop sown, availability of field (ready for sowing), the financial condition of the farmer, etc. As a result, even the crops that are sown in the same season differ in time, meaning, they have different sowing times. Due to this all fields having the same crop cannot be mapped at one time. Also in a particular season, a different variety of crops are cultivated in a given area due to which there is spectral overlap between the crop classes when observed through remote sensing images. This spectral overlap between the target crop class and other crops in the area needs to be reduced as they possess the potential to be classified as the target crop.

Hence, a multi-temporal approach was adopted which provides phenology as changes occurring within crop information which is unique and helps in reducing the spectral overlap between the crop classes. A well-known technique for finding out optimum features is separability analysis which was performed to find the optimum dates with minimum spectral overlap between the crop classes.

1.3. Research Objectives

The main objective of this research was to study and design a CNN-based deep learning algorithm for specific crop mapping using temporal data sets.

Sub-objectives:

The research focuses on developing a CNN and an integrated CNN-LSTM model focusing on crop type classification. Working of CNN and the integrated version of the algorithm was studied for mapping a single crop class. Heterogeneity within a class was also studied and was aimed to minimize using the fuzzy algorithm. Instead of taking the global mean of all the input samples, it has been experimented to use individual sample as a mean in the fuzzy based classifier which can handle heterogeneity. The specific objectives are stated below:

1. Finding the minimum number of temporal images required to map specific crops using multi-temporal data.
2. To study and optimize the CNN and integrated CNN-LSTM model while monitoring parameters like no. of iterations (epoch), no. of layers, model loss and accuracy, no. of neurons, learning rate, etc. for specific crop mapping.
3. To study heterogeneity (intra-class variability) within a class through deep learning-based classifiers and compare it with fuzzy Modified Possibility c-Mean (MPCM) classifiers.
4. Testing dual-sensor data approach for classification.

1.4. Research Questions

The following research questions will be addressed:

1. Referring to sub-objective 1:
    a. What is the minimum number of temporal images required to map a specific crop while working with temporal data?
2. Referring to sub-objective 2:
    a. What are the values of the hyper-parameters that produce the best results?

   b. How much is the model loss and final accuracy for which the proposed model gives the best result for specific crop mapping?
3. Referring to sub-objective 3:
   a. Which method best handles the heterogeneity within a crop class – deep learning model or MPCM approach?
4. Referring to sub-objective 4:
   a. Does the SAR data give any extra information as compared to the optical dataset?
   b. Which approach is better - single sensor or dual-sensor approach for classification?

1.5. Innovation Aimed at

The innovation in this research was focused on how CNN and integrated CNN-LSTM models can be applied for temporal data processing for specific crop mapping while handling heterogeneity within a class, with minimum hardware resources.

1.6. Research Approach

In this research, Sentinel-2 time-series datasets were used for the classification of specific crop types and to achieve the objectives. The research was conducted in two different time phases, the first phase between April 2020 to June 2020 where the identified class of interest (target classes) were 'Maize (Corn)', 'Mentha (Mint)' and 'Guava' while the second phase between October 2020 to March 2021 where the classes of interest (target classes) identified was 'Early Potato', 'Mid-Potato' and 'Late Potato'. The phase-1 aims at level-2 classification as the classification was done for distinct crop classes identified in the study area, while the phase-2 aims at level-3 classification as the classification was done within a single crop (Potato) where the target crops were divided depending upon the sowing time of the crop. During Phase 1, the Landsat-8 dataset was incorporated to fill in the temporal gaps due to cloud cover. The data set was pre-processed and resampled to bring the spatial resolution of the Landsat-8 dataset equal to the Sentinel-2 dataset, i.e., 10m. Class-based Sensor Independent Normalized Differential Vegetation Index (CBSI-NDVI) was calculated for the dataset considering one target class at a time. This data was further utilized to perform separability analysis which gave the optimum number of dates minimizing the spectral overlap between the classes, necessary for specific crop classification. Using optimum date datasets, training data was generated to be fed as an input to the Convolutional Neural Network (CNN) model. The classified outputs generated from the models was assessed using precision, recall and f1-score. However due to the limited training data available, this evaluation alone could not be considered for assessing the performance of the classifier. Therefore, the classification results obtained were also assessed using Mean Membership Difference (MMD) (Singh, Kumar, & Upadhyay, 2020) evaluated between interest class and non-interest class, and the variance within the target crop field. The performance of the model was assessed graphically by observing the model accuracy and loss for each target crop class. Also, the hard classification outputs were generated and the target crop class field was observed to assess the performance of the classifier. Following a similar process, results were obtained for the hybrid deep learning model, i.e., Convolutional Neural Network – Long Short Term Memory (CNN-LSTM) model and compared with the CNN model. The fuzzy Modified Probabilistic c-mean (MPCM) classifier was used to handle the heterogeneity within a crop class (intra-class variability) and the output was compared with deep learning classifiers' results. This classifier was adopted due to its ability to classify non-overlapping clusters and handle noise in the dataset. Lastly, the dual-sensor approach was tested by replacing one of the Sentinel-2 optical images with a Sentinel-1 SAR image to investigate an additional information added to the data, if any.

1.7. Structure of the Thesis

This thesis has been divided into six chapters. The first chapter 'Introduction' talks about the motivation behind the research work and how this research work has been carried out. It also includes the research objectives and corresponding research questions along with the innovation introduced in the work. The second chapter 'Literature Review' includes the details of the previous research work conducted for similar application along with the conceptual explanation of the methodologies and technologies used in this research. The third chapter 'Study Area and Datasets used' explains in detail the study area and datasets that were chosen for this research and gives the motivation to choose them. It also discusses in detail how various crop classes were divided and used in this research. The methodology of this research is explained in detail in the fourth chapter as 'Research Methodology'. The results obtained were visualized and discussed alongside in the fifth chapter 'Results and Discussions'. The conclusions drawn from the analysis of the research along with future recommendations were presented in the last chapter 'Conclusions and Recommendations'.

2. LITERATURE REVIEW

The interaction of the Remote Sensing (RS) and Geo-information field have a lot of potential and significance in the use and analysis of spatial data. Geographic Information Systems (GIS) is a broader field that extends remote sensing fields to disciplines like photogrammetry, earth sciences, computer science, cartography, etc. (Courville, 1987). While classifying crop classes, the spectral profile of a crop plays a vital role as each crop has a varying and unique spectral reflectance. But in some cases, the spectral reflectance of two crop classes are very similar to each other which makes it difficult to classify. For this reason, the phenology of the crop is studied through time-series data that minimizes the spectral overlap between the classes (Bargiel, 2017). Time-series data extracts seasonal changes in the spectral profile of a crop type that is unique to each crop.

Every year in India, state-wise yield estimation is calculated for each crop. This is performed to estimate the area under cultivation for each crop which further could provide a rough figure of crop yield in that season. Therefore, specific crop mapping plays an important role in estimating the area under cultivation for each crop.

2.1. Role of Vegetation Indices in Time Series

A lot of studies have suggested that temporal features are more useful than spatial features during classification (Simonneaux et al., 2008). The datasets like Sentinel-2, which are freely available with a high temporal resolution of five days and spatial resolution of around 10m have created the possibility to use the data for research in the applications like crop classification (Mazzia et al., 2020). Apart from the temporal feature extraction, other factors that make the selection methods more difficult are designing a model and extracting features, inter-class variability and uncertain atmospheric conditions.

When the sunlight hits the ground surface, it absorbs some of the energy while reflecting the remaining back into the atmosphere. The pigments in the plants absorb the visible light and the shape of the leaves determines how much of the energy is reflected into the atmosphere. Near-infrared (NIR) wavelength is strongly reflected into the atmosphere by the vegetation dependent on leaf structure (Paul Przyborski, 2000). In the field of remote sensing, the Normalized Differential Vegetation Index (NDVI) is the most commonly used vegetation index for multi-temporal studies (Zhang et al., 2018). NDVI is the ratio of the difference between canopy reflectance of NIR and red bands as shown in Equation (1).

$$\text{NDVI} = \frac{NIR-Red}{NIR+Red} \tag{1}$$

2.1.1. CBSI-NDVI Index and its importance

With time, new satellites have been launched with an increased number of spectral bands as compared to previous satellites like Landsat. So, a new Class-Based Sensor Independent Normalized Differential Vegetation Index (CBSI-NDVI) was introduced which took into account all the bands that are available to produce the result (Kumar, 2012). Equation (2) represents the calculation of CBSI-NDVI:

$$\text{CBSI-NDVI} = \frac{Max - Min}{Max+Min} \tag{2}$$

Here min and max represent the bands in the image with minimum and maximum Digital Number (DN) values respectively. The selection of the minimum and maximum band depends on the target crop. The spectral signatures for each band are observed at the field of the target crop. The band which had maximum spectral reflectance values at the target crop field is chosen as the 'maximum' while the band that had minimum spectral reflectance values at the target crop field is chosen as the 'minimum'. In this way, minimum and maximum bands are computed for each date image to calculate the CBSI-NDVI index. This index has proven to enhance the spectral information of the image and creates better separation between the crop classes.

2.2. Deep Learning based classifier

Artificial Intelligence (AI) is a technology that enables computers to mimic human behaviour (Simmons & Chappell, 1988). AI has its applications in fields like robotics, speech recognition, computer vision, etc. Machine learning (ML) is a sub-domain of AI that can learn without explicitly being programmed and has its applications in logistics, regression, etc. (Ian Goodfellow et al., 2016) Representation learning is a more specific domain under ML where the model learns from the input data provided. Deep Learning (DL) is a type of representation learning that often extracts hidden patterns from the data using neural networks. The relationship between these domains can be seen in figure 1.

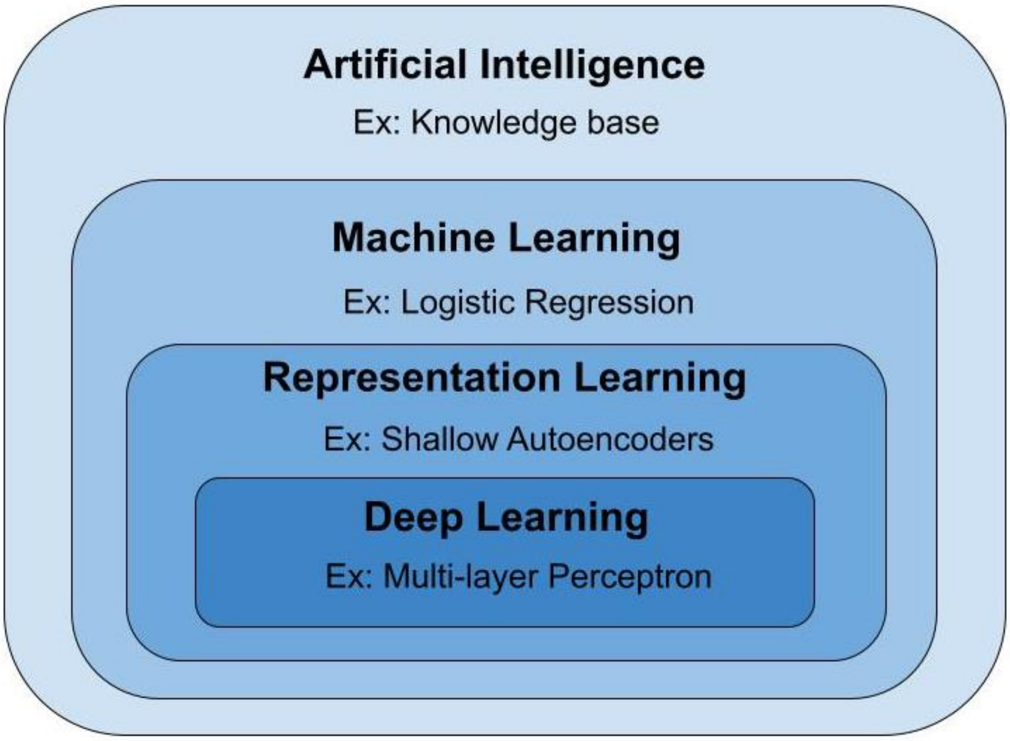


Figure 1: Deep learning is a type of representation learning

Classification using machine learning techniques is being practiced and has proven to produce improved results as compared to conventional parametric techniques. The training set that is being fed to the model needs to possess certain aspects. It is very important that there is no or minimum imbalance between the training samples and the features for the classifier to perform at its best (Z. Liu & Sun, 2008).

2.2.1. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN), also known as ConvNets, were introduced in the 1980s by YannLeCun (Lecun et. al., 1998), who was a postdoctoral computer science researcher. It utilized the ideas of a Japanese scientist named Kunihiko Fukushima (Fukushima, 1988) who invented a basic image recognition neural network.

CNN has been used for image classification since the 1990s (Choi et al., 2005). But due to the hardware constraints and the inability to perform complex computations, this area was not explored much previously. Figure 2 depicts the structure of a multi-layer perceptron (MLP) which is a class of feed-forward artificial neural networks. Here, each circle represents a neuron. A neuron in a computational unit that receives input from the input wires and gives output after performing some computation. Layer 1 is known as the input layer and contains the input given to the model (Nielsen, 2015). Layer 2 is known as the hidden layers and contains an activation function that performs computation and produces an output which is further given to the output layer, i.e. Layer 3. Each neuron is a matrix of values and each wire connecting the neuron bears some weight which is used to evaluate the neuron in the next layer. There can be any number of hidden layers in a neural network. Also, a bias is introduced in every layer. So, output for is computed as mentioned in Equation (3):

Output = sum (weights*inputs) + bias (3)

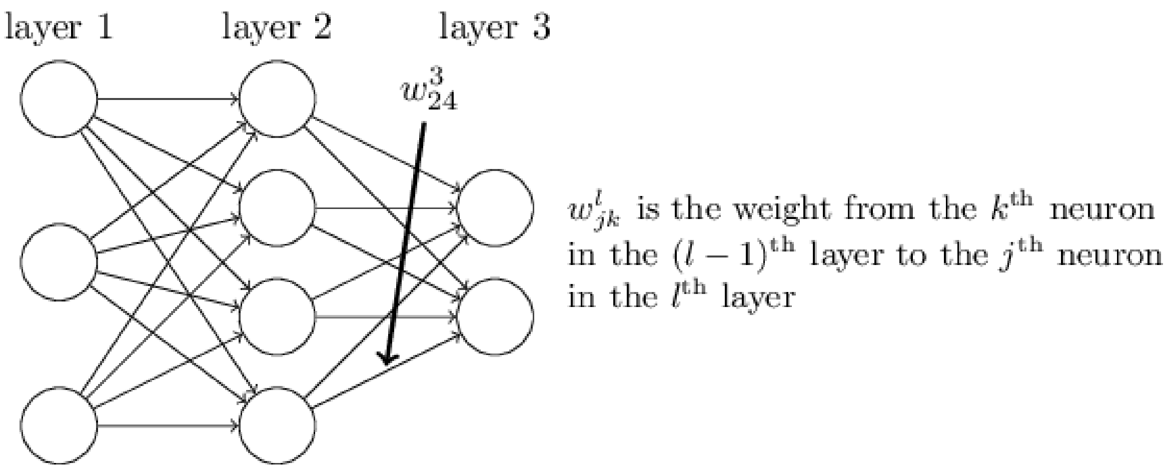


Figure 2: Structure of a Multi-layer Perceptron (Nielsen, 2015)

CNN can be considered a special case of MLP where one input produces a single output after going through a number of layers. There are various layers constituting CNN as can be seen in its architectures, as shown in figure 3, which can increase or decrease based on the application for which it is being developed (C. Zhang et al., 2018). The hyper-parameters play a key role in determining the structure of the model and also decide how the model will be trained on the data. Some of the hyper-parameters include a number of hidden layers, dropout layer, activation functions, learning rates, number of epochs, and, the batch size (Loussaief & Abdelkrim, 2018).

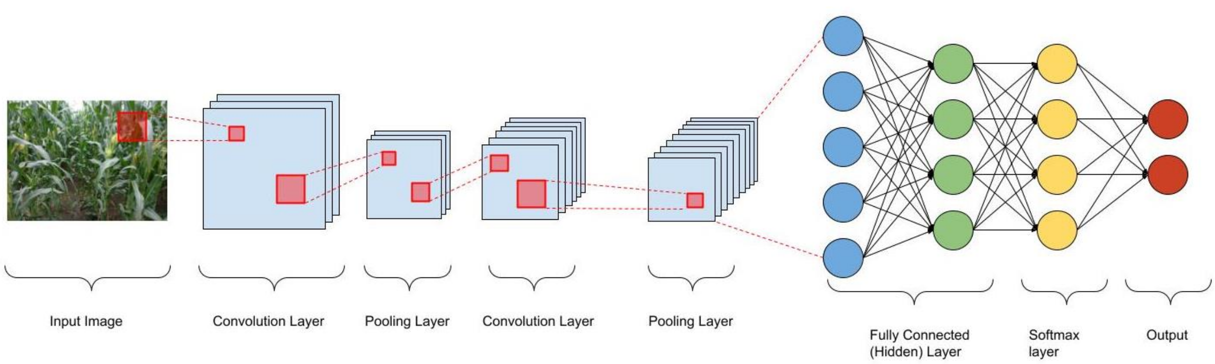


Figure 3: Connectivity of the layers in a CNN model

2.2.1.1. Convolution Layer

The first layer is always the convolution layer. In this layer, convolutions (or filters or kernels) are applied to the input layer in such a way that it sweeps over the whole image. The dimension of the kernel applied varies based on the application but its size is always smaller than the input image. One of the main advantages of using a kernel is that it works with local perception information (Chunjing et al., 2017). This helps the kernel to learn from a specific portion of the input image by increasing the dimensionality of the input. When the first kernel is applied, weights for each kernel are initialized by learning and these weights remain the same for all the kernels applied in the convolution. In this way, parameter sharing takes place (J. Yang & Li, 2017).

2.2.1.2. Activation Function

The output from a neuron is fed directly into an activation function which introduces non-linearity to the data (Sharma, Sharma, & Anidhya, 2017). This is a crucial step to extract the hidden information from the input image. The most commonly used activation function is Rectified Linear Unit (ReLU) activation function due to its ability to performing faster computations (Gonzalez, 2007) as compared to other activation functions such as tanh, sigmoid, leaky ReLU, etc.

2.2.1.3. Pooling Layers

Since multiple convolutions are applied to the input data, the dimensionality of the data increases. To reduce the dimensionality of the output from the previous layer, a pooling layer is added in the model which also, in turn, saves the computational resources (Chunjing et al., 2017). When a model includes a large number of layers in it, there are high chances of overfitting the model. So, a dropout layer is included in the CNN architecture to prevent the model from overfitting (Mele & Altarelli, 1993).

2.2.1.4. Fully connected layers

In the end, a fully connected layer or dense layer is included which predicts the class of the output obtained. This uses a softmax activation function to predict the class (Ian Goodfellow et al., 2016).

2.2.2. Recurrent Neural Networks (RNN)

An advantage of using CNN is that the algorithm down-samples the input image without losing its original characteristics. Hence, it requires less computational resources and is more energy-efficient. With the advancement in technology, recently there has been more exploitation of CNN and its hybrid with RNN-LSTM algorithm in the field of remote sensing (Zhong et al., 2019). CNN has proved to be beneficial in extracting the spatial features from an image while LSTM brings out the phonological characteristics which are useful in specific crop mapping (Sun et al., 2019). RNN is a feed-forward neural network that stores the output of the previous computation in its memory as shown in figure 4.

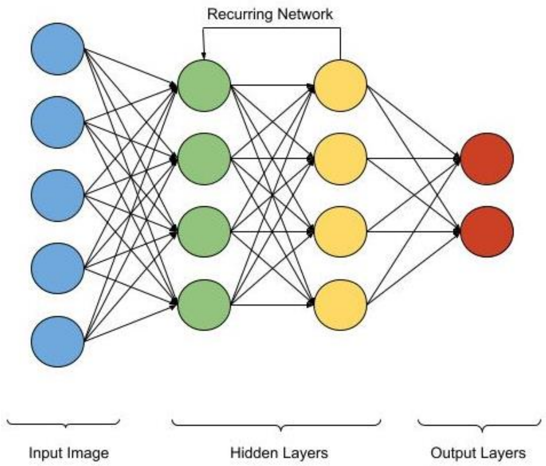


Figure 4: RNN Unit

So, the network considers what it learns from the previous output along with the current input for making a decision (Wu & Prasad, 2017). Although after some iterations in RNN, some values seem to vanish from the memory when they become less dependent, also known as the vanishing Gradient problem. Therefore, a special case of RNN called LSTM (Long-Short Term Memory) is applied where it stores the information in its memory for a longer period. This helps in improving the predicting abilities of the model as the information about the neighbourhood of a pixel is stored in the memory.

Studies suggest that deep learning methods can be exploited by incorporating them with neural networks to optimally extract the temporal information along with the spatial and spectral information (Courville, 2017). In some applications, 2D CNN has been used to extract spatial data from high-resolution images (Maggiori et al., 2017). But in the applications where there is limited availability of training data, 1D CNN and its hybrid approach have proved to be more effective (Kiranyaz et al., 2019). An advantage of adopting a hybrid approach is that the research work can be tested on CPU based hardware (Wang et al., 2017) making the proposed model energy-efficient.

2.2.3. Deep learning for crop type mapping using Remote Sensing

The classified maps produced by utilizing the RS data plays a crucial role in agricultural crop studies. The deep learning algorithms have been extensively utilized for crop mapping over the time and have proved to produce reliable results for complex and heterogeneous datasets (B. Yang & Xu, 2021). With the free availability of medium-high resolution RS dataset such as from Sentinel-2 mission, the performance of the classification outputs generated from the deep learning classifiers have improved significantly (Garnot, et. al., 2019). The deep learning classifiers automate the process of classification which in turn reduces both time and manual effort (Ashourloo et al., 2020). Various hybrid deep learning models have also been studied for crop classification and have proved to give better performance. Although, one of the main challenge during crop classification is the heterogeneity within a class due to non-uniform dataset (Turkoglu et al., 2021).

2.3. Fuzzy based classifier

The conventional classifiers that used set theory for classification have been incapable of producing adequate classification results due to their incapability of handling the mixed pixels (Murmu & Biswas, 2015). These classifiers classified a given pixel into a particular class or label. Therefore, soft classification techniques used by fuzzy classifiers have been applied where membership values are assigned to a pixel. The membership value is the degree of belongingness of a pixel to a particular class or label (Foody, et al., 1992). In this technique, a pixel can belong to more than one class/label at a given time resulting in unique pixel values.

The key advantage of using a fuzzy classifier in remote sensing for classification is that it takes into consideration the real-world situation. Meaning, instead of considering all the boundaries as crisp, it assigns more than one class/label to a pixel in the form of membership values. These values lie between 0 and 1 (Foody et al., 1992).

Some of the most widely used fuzzy-based classifiers are the Fuzzy c-Mean (FCM) classifier and Possibilistic c-Mean (PCM) classifier with both of them having their advantages and disadvantages. The FCM classifier has a major assumption that the sum of the membership values of all the classes that a given pixel is composed of should always be equal to 1. FCM is also sensitive to the outliers which result in the formation of unrealistic clusters. The PCM classifier could be seen as an improved version of FCM as it handles the outliers more effectively, requiring good initialization. But the major drawback of the PCM classifier is that it results in coinciding clusters. This feature of PCM is undesirable when the aim is for single-class classification (Misra et al., 2014).

2.3.1. Modified Possibilistic c-Means Classifier (MPCM)

The specific crop classification is a challenging task due to the overlapping nature of the spectral reflectance of each crop type. This becomes even a major obstacle when performing a higher level of classification like level-2 and level-3 classification. The level-2 classification includes the division of study area into distinct crop types while level-3 classification includes the division of study area into classes within a crop type.

The MPCM classifier is an enhanced version of the PCM classifier. It overcomes the drawbacks of the PCM classifier which includes its ability to classify the dataset into non-overlapping clusters (Singh et al., 2020). The working of the MPCM classifier is explained in detail in section 4.6.

2.4. Dual-sensor Approach

Over time, optical datasets have been utilized exclusively for land use land cover (LULC) classification. The integrated dataset from different optical satellites have largely been utilized due to its compatibility in the spectral domain. However, utilization of the optical datasets is restricted during cloud cover leading to capture of low spectral reflectance (Stroppiana et al., 2015). The cloud cover over a region in turn also results in low spectral separability between different classes. Unlike the optical datasets like Landsat-8 or Sentinel-2 imagery, SAR datasets are independent of the spectral domain and instead use the structural and dielectric properties of the target (S. Liu et. al., 2019). There have been studies where the classification accuracy has significantly increased using the integrated optical and SAR datasets in the MLP classifier (Kussul et al., 2016). Hence, the dual-sensor approach is studied in this research to test its performance with the deep learning classifiers.

3. STUDY AREA AND DATASET USED

3.1. Study Area

The study area chosen for this research work lies in the Ludhiana district of Punjab state in India as shown in Figure 5. Punjab is an agricultural state in the Northern part of India. The Ludhiana district is located at 30.9010°N and 75.8573°E alongside the river Sutlej. The major crops grown in this region are paddy (rice), wheat, and maize (corn). The research was carried out in two phases: Phase-1 covering the period from April 2020 to June 2020 while Phase-2 covering the period from October 2020 to March 2021. The study area remained the same for both the phases but the target or interest crop varied depending upon the period.

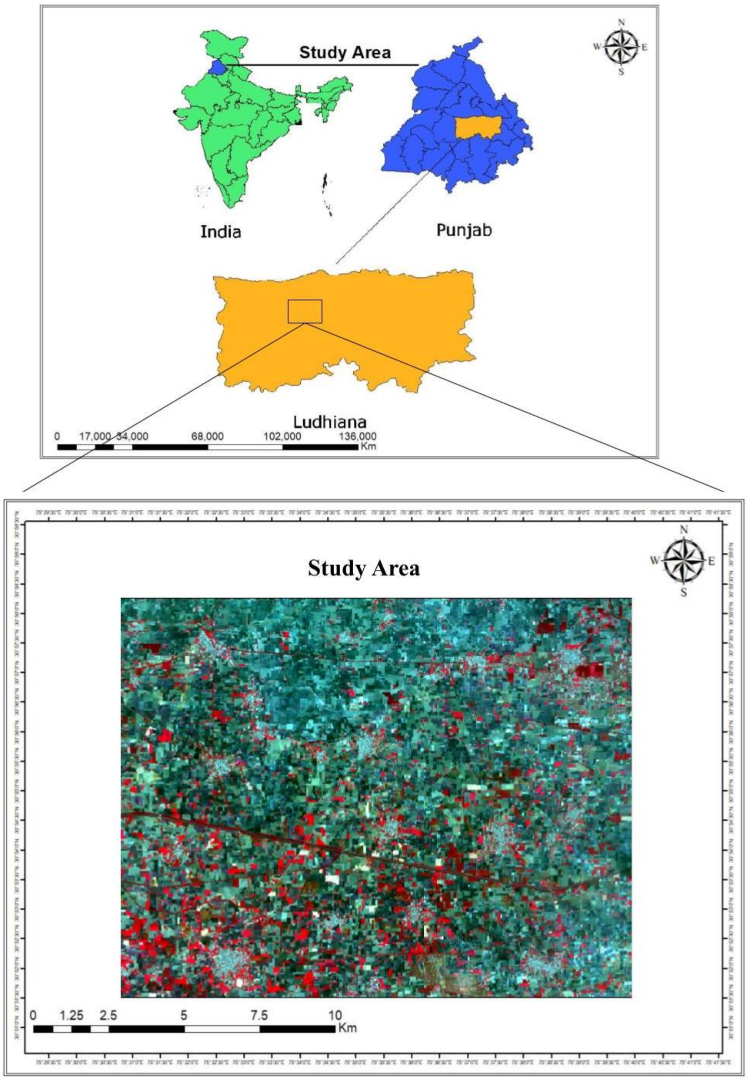


Figure 5: Study Area

For phase-1, the target crops were maize (scientific name - Zea Mays), mentha (also known as mint) and guava (scientific name - Psidium guajava) orchard. All the target crops have different sowing and cultivation periods which results in discrete phenology and hence discrete spectral signatures at a given instance.
The initial growth stages of the Maize crop were observed in the study as depicted in table 2. The duration of cultivation of the crop in the field is around 90 days (Jaskaran & Simerjeet, 2020). During the initially observed period, a previous crop, i.e., the wheat crop was present in the field which was harvested in the

month of April. Subsequently, the field was prepared for the sowing of the maize crop and the maize was sown in the second fortnight of May. In the month of June mainly two stages of the crop were observed which were the seedling stage, where the crop starts to emerge from the ground, and the knee-high stage, where the height of the crop reaches the knee of a person.

Table 2: Maize crop growth stages covered

Time Period	**Crop Stage of Maize**	**Dates covered in dataset**
First Fortnight of April	Previous crop (Wheat)	02-Apr-2020
Second Fortnight of April	Crop Harvested (Fallow field)	22-Apr-2020
First fortnight of May	Current fallow	07, 12-May-2020
Second fortnight of May	Preparation of field for Maize crop and pre-sowing irrigation	27-May-2020
First fortnight of June	Sowing of Kharif Maize and seedling stage	15-Jun-2020
Second fortnight of June	Knee-high stage	21-Jun-2020

The later stages of the growth cycle of the Mentha crop were observed in the research as depicted in table 3. The duration of the crop in the field is around 140-150 days (Jaskaran & Simerjeet, 2020). The crop was sown in the second fortnight of February month. Due to the lower temperatures, the initial growth of the crop is slow. Hence in April month, even after the emergence of the crop above the ground surface, the coverage of the crop is minimal. Due to this, the spectral signature observed through the satellite images seems to be similar to a fallow field. In the second fortnight of April, the temperature starts to rise, and the crop starts to cover the ground, hence vegetative crop signatures were observed in the satellite imagery. The crop starts to grow and spread throughout the field giving more intense spectral signatures. In the first fortnight of June, the crop reaches the Senescence stage, meaning the crop starts to mature making it ready for harvest (Sukhmal et. al., 2004). By the second fortnight of June, the crop was harvested and the field again gave signatures of the fallow field.

Table 3: Mentha crop growth stages covered

Time Period	**Crop stage of Mentha**	**Dates covered in dataset**
Second Fortnight of February	Planting of Crop	-
First fortnight of March	Emergence	-
Second Fortnight of March	Emergence	-
First Fortnight of April	Emergence of crop	02-Apr-2020
Second Fortnight of April	Vegetative growth (partially covering ground)	22-Apr-2020
First fortnight of May	Vegetative growth (Fully covering the ground)	07, 12-May-2020
Second fortnight of May	Continued vegetation growth	27-May-2020
First fortnight of June	Senescence	15-Jun-2020
Second fortnight of June	Harvesting of crop	21-Jun-2020

The third target crop in phase-1 was Guava orchards which is an evergreen crop. In India, Guava is grown in tropical and sub-tropical climatic conditions owing to its wider adaptability (Surinder & Jaskaran, 2020). The guava orchards observed in this research were around 10-12 years old. It bears fruits twice a year, i.e., during rainy and winter seasons. In this research, the fruit-bearing period of the crop is covered. The flowering in the plant starts around the second fortnight of April and the fruits start to grow around the second fortnight of May.

Another crop found in the study area during the same period was bajra (scientific name – Pennisetum glaucum). The bajra crop was considered as the non-interest class in this phase because the size of the crop field was very small as compared to the other crops in the area resulting in less spectral information.

For phase-2, the target crop was potato (scientific name - Solanum tuberosum). The level-3 classification was adopted where the potato crop was sub-divided into three target crop classes namely early potato, mid potato and late potato based on the sowing period of the crop as shown in table 4.

Table 4: Phase 2 target crop classes' growth considered

Sentinel - 2	**Early Potato**	**Mid Potato**	**Late Potato**	**Dates covered in Dataset**
Second fortnight of October	Sprouting & Emergence	-	-	14, 24-Oct-2020
First Fortnight of November	Foliage	Sprouting & Emergence	-	03, 08, 13-Nov-2020
Second Fortnight of November	Foliage	Foliage	Sprouting & Emergence	28-Nov-2020
First Fortnight of December	Tuber Formation	Foliage	Foliage	03-Dec-2020
Second Fortnight of December	Tuber Development	Tuber Formation	Foliage	18, 23-Dec-2020
First Fortnight of January	Senescence	Tuber Development	Tuber Formation	-
Second Fortnight of January	Early Harvesting	Senescence	Tuber Development	27-Jan-2021
First Fortnight of February	Late Harvesting	Early Harvesting	Senescence	01, 06-Feb-2021
Second Fortnight of February	Field preparation and sowing of next crop	Late Harvesting	Early Harvesting	21, 26-Feb-2021
First Fortnight of March	-	Field preparation and sowing of next crop	Late Harvesting	08-Mar-20201

This is called level-3 classification because the classification was done within a particular crop class. The target crop were selected based on the period of sowing of the potato crop in the field where early potato, mid-potato and late potato crops were sown in second fortnight of October, first fortnight of November and second fortnight of November respectively. Hence, different phenological stages were observed for the target crops on a particular date as shown in table 4. The duration of early potato, mid potato and late potato crops in the field is around 90-100 days, 90-100 days, and 100-110 days respectively. There can be slight variation in the duration of the crop in the field depending upon the variety sown (Surinder & Jaskaran, 2020b). The Early potato crop was sown in the second fortnight of October, the mid potato crop was sown in the first fortnight of November and the late potato was sown in the second fortnight of November. Phase-2 covered the whole phenological cycle of the target potato crops including the sowing and the harvesting of the crop. The crop growth starts with sprouting and emerging of the crop where the seed starts to germinate and the seedling start emerging from the ground. During the foliage stage, the leaves starts to emerge from the plant. An auxiliary bud starts to form on the potato stem during the tuber formation and the potato starts growing in size during the tuber development stage (KG, 2021). The Senescence stage indicated that the crop is proceeding towards maturity. In the end, early harvesting takes place which includes cutting and collection of the fully grown tuber mainly. Also, late harvesting of crops takes place where fully grown tubers are harvested at a slightly later stage. The non-interest classes considered in the study area for this period were guava, wheat as 'other crop' and the settlement.

3.2. Dataset Used

3.2.1. Phase 1

This phase includes the crop types which are cultivated between April 2020 and June 2020. The dataset used in this phase was the freely available Sentinel-2 optical dataset. The Sentinel-2A and 2B satellites are a part of the Copernicus Program launched by the European Space Agency (ESA) in June 2015 and March 2017 respectively (Sentinel missions overview, 2021). Sentinel-2A and 2B satellites carry optoelectronic multispectral sensors for which the resolution varies from 10m in the visible region to 20m in the near-infrared (NIR) to 60m in short-wave infrared (SWIR) - Cirrus spectral zones. The main focus of the Sentinel-2 mission was land monitoring which included monitoring vegetation, soil and coastal areas (User Guides - Sentinel-2 MSI - Overview, 2021). This mission comprises of two polar-orbiting satellites providing medium-resolution optical imagery. These satellites were equipped with a Multi-Spectral Instrument (MSI) sensor that measures Earth's reflected radiance in 13 spectral bands. The sentinel bands that were used in this research are as mentioned in table 5:

Table 5: Sentinel-2 bands details

Band	Resolution
Band 2 – Blue	10m
Band 3 – Green	10m
Band 4 – Red	10m
Band 5 - Vegetation red edge	20m
Band 6 - Vegetation red edge	20m
Band 7 - Vegetation red edge	20 m
Band 8 – NIR	10m
Band 8a - Narrow NIR	20m
Band 11 – SWIR	20m
Band 12 – SWIR	20m

In June, cloud cover was observed over the study area in the Sentinel dataset which resulted in the temporal gap. To fill this temporal gap, Landsat 8 dataset was also used and integrated with the Sentinel dataset. Landsat 8 mission was collaboratively developed by NASA and US Geological Survey (USGS) and was launched in February 2013 (Survey, 2016). The Landsat 8 satellite consists of two sensors – Operational

Land Imager (OLI) and Thermal Infrared Sensor (Survey, 2018) which provides global coverage with the spatial resolution ranging from 15 meters for Panchromatic bands to 30 meters for visible, NIR & SWIR bands to 100 meters for thermal bands. The Landsat bands used in this research are as mentioned in table 6:

Table 6: Landsat-8 band details

Band	**Resolution**
Band 1 - Coastal Aerosol	30m
Band 2 - Blue	30m
Band 3 - Green	30m
Band 4 - Red	30m
Band 5 - NIR	30m
Band 6 - SWIR-1	30 m
Band 7 - SWIR-2	30m
Band 8 - Panchromatic	15m
Band 9 – Cirrus	30m
Band 10 - TIRS-1	100m
Band 11 - TIRS-2	100m

3.2.2. Phase 2

This phase focuses on the potato crop cultivated in the given study area between the period of October 2020 and March 2021. This phase used Sentinel - 2A and 2B datasets using the optical bands as shown in table 5 above.

The Sentinel-1 dataset was used to test the dual-sensor approach and test if using Synthetic Aperture Radar (SAR) data along with optical data provides any essential information for classifying a specific crop. Sentinel-1 was the first satellite launched under the Copernicus Program by ESA (Sentinel-1 Instrument Payload, 2021). It also consists of twin satellites, i.e., Sentinel – 1A and Sentinel – 1B. It had a swath of 80 Km and a spatial resolution of 5m * 5m. It works in dual-polarization mode, i.e., VV+VH and HH+HV where the first symbol represents transmission (V for vertically and H for horizontally) while the second symbol represents receiving (V for vertically and H for horizontally). Sentinel- 1 produces two data products, Ground Range Detected (GRD) and Single Look Complex (SLC). GRD dataset is used in this research work.

3.3. Data Availability

In this research, the multi-temporal optical dataset was used. It included both Sentinel - 2A & 2B and Landsat-8 datasets for the phase-1 study. For testing the dual-sensor approach for phase 2 crops, Sentinel - 1A & 1B SAR dataset was also included. The datasets used in the research are as shown in table 7 and table 8:

Table 7: Temporal dataset for Phase 1 crops

Sentinel - 2	Landsat - 8
02 April 2020	
22 April 2020	
07 May 2020	
12 May 2020	
27 May 2020	
	15 June 2020
21 June 2020	

Table 8: Temporal dataset for Phase 2 crops

Sentinel - 2	Sentinel - 1
14 Oct 2020	
24 Oct 2020	
03 Nov 2020	
08 Nov 2020	
13 Nov 2020	
28 Nov 2020	
03 Dec 2020	
18 Dec 2020	
23 Dec 2020	
27 Jan 2021	27 Jan 2021
01 Feb 2021	
06 Feb 2021	
21 Feb 2021	
26 Feb 2021	
08 Mar 2021	

3.4. Field Data Collection

The field visit was conducted in the study area which included visiting the agricultural fields in the villages lying in the west part of the Ludhiana district as shown in figure 6 for phase-1 crops and figure 7 for phase-2 crops. The geo-location of the fields was collected along with the type of crop sown in that field. The geo-locations were collected using the 'mobile topographer' app which gave high accuracy of around 1-2m. The sowing date of the crops was also confirmed by the farmers owing the land. The data collection for phase-1 was done on field visit conducted in the third week of July 2020. The data collection for phase-2 was done once in November 2020 and then later in February 2021.


(a) Maize


(b) Bajra


(c) Mentha


(d) Interaction with farmers


(e) Guava Orchards

Figure 6: Field visit photographs of phase -1


(a) Early Potato field


(b) Mid potato field


(c) Late potato field

Figure 7: Field visit photographs of phase -2 as observed in November, 2020

4. RESEARCH METHODOLOGY

4.1. Methodology Adopted

Figure 8 indicated the flowchart of the methodology that was adopted in this research. The multi-temporal data was used to classify the crop types by studying the phonological changes occurring over time. The research was divided into two phases capturing distinct seasons and crops while maintaining the spatial extent of the study area. The field visit was conducted in both phases where the geographical location of the field was collected along with the name and time of sowing of the crop sown in that particular field. For phase-1, three crops were selected as the target crop which was maize (Corn), Mentha (Mint), and Guava orchard. For phase-2, the potato crop was divided into three sub-classes which were early potato, mid potato and late potato. The pre-processing of the dataset was carried out as per the requirements to integrate distinct datasets. Thereafter, CBSI-NDVI was computed to better extract the spectral information of the crop from the dataset available. This helped by enhancing the spectral signatures for are our target crop classes in the study area and decreasing the spectral signature of non-interest classes. While the spatial extent of the study area was maintained, the spectral domain was minimized in a way that it extracted most of the spectral information required from the data. Hence, separability analysis was carried out in which the optimum number of dates were extracted as the output for each crop class that gave maximum separation of the target/interest class from the non-interest classes. These dates were then considered in the final dataset that was fed as the input to the deep learning models.

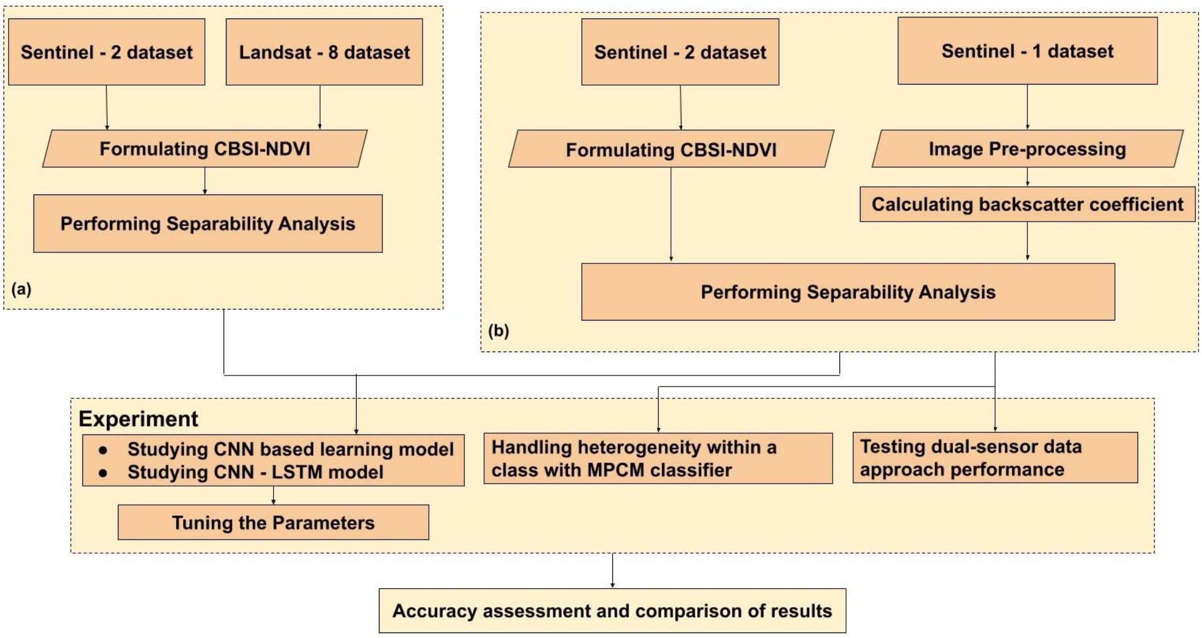


Figure 8: Methodology Flowchart for (a) Phase 1, and (b) Phase 2

The CNN model was generated and optimized for all the target crop classes in both phases. Subsequently, the classified results were produced from the model. Similarly, the CNN-LSTM model was further explored and optimized with the same dataset and compared with CNN model outputs. The classified results were produced in the form of both soft and hard outputs. These classified outputs from CNN and CNN-LSTM deep learning models were analysed using f1-score, precision and recall. Apart from this, the accuracy assessment was carried through Mean Membership Difference (MMD) and Variance. The performance of the deep learning models was also taken into consideration while evaluating the outputs.

Also, the fuzzy-based classifier was trained using the same training data and tested for heterogeneity within a crop class. The heterogeneity within a crop class was analysed and compared between the outputs generated from the fuzzy classifier and deep learning classifier. This was done to test which classifier produces better results even when there is heterogeneity (intra-class variability) within a crop class.
Lastly, the dual-sensor approach was tested where one Synthetic Aperture Radar (SAR) date image from Sentinel – 1 was incorporated with the training data by replacing one of the Sentinel – 2 images. This was done to test if the SAR dataset adds any further information to the dataset which could subsequently help in aiding better classification results.

4.2. Pre-Processing

When working with Remote Sensing data, pre-processing acts as a very crucial step as this removes unwanted noise due to atmospheric conditions, making the dataset more interpretable. This becomes highly important when integrating datasets from different satellites as the final data needs to be compatible with each other in both spectral and spatial terms. All the fields in the dataset need to coincide especially when working with smaller fields as even a slight error in this phase could be transferred to the following processing steps.
The Sentinel-2 optical dataset utilized in this research was level-2 atmospherically corrected data. This dataset was freely accessible and downloaded from the Copernicus Open access hub made available via the European Space Agency (ESA). Also, the Landsat-8 dataset was utilized in the study which is also level-2 atmospherically corrected data and made freely accessible from USGS Earth Explorer. Both these datasets are georeferenced in the UTM projection system but available at different spatial resolutions. The Sentinel-1 dataset was present at 10m resolution while the Landsat-8 dataset was present at 30m resolution. Therefore, the Landsat-8 image was resampled at 10m spatial resolution to make it compatible to use with the Sentinel-2 dataset.
For the Sentinel – 1 dataset which was used for testing the dual-sensor approach, pre-processing was carried out in Sentinel Application Platform (SNAP) which includes orbital correction followed by radiometric calibration as shown in figure 9. Then multi-looking was performed to make the pixels square. Then speckle filtering was performed which removes unwanted noise from the dataset. Then backscatter coefficient was calculated which produces results in decibels (dB). Finally, to bring data similar to that of optical data, the dataset was exported as geo-tiff. The upscaling of the dataset was performed in ERDAS Imagine software to finally convert it into unsigned 8-bit. This was necessary as the Sentinel-2 optical data was present at an 8-bit scale. Therefore Sentinel-1 dataset was made compatible with the Sentinel-2 dataset to be utilized further.

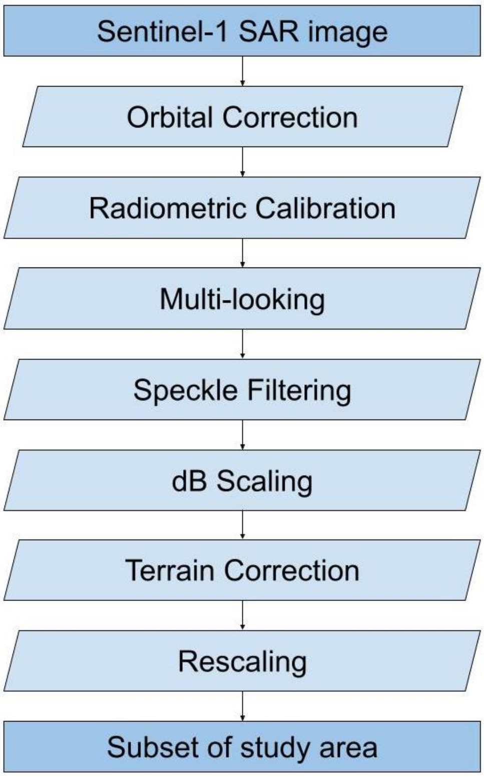


Figure 9: Pre-processing for SAR data

4.3. Calculating Temporal Index

After the data was pre-processed, the vegetation index was calculated for all the optical datasets used in both phases. The vegetation indices play an important role in enhancing the spectral signature of the vegetation in the study area by distinguishing it from other non-interest classes. In this research, CBSI-NDVI was utilized which is a slightly modified version of the NDVI index. While evaluating the CBSI-NDVI index for a target crop class, the band that gives maximum reflectance was considered as NIR while the band that gives the minimum reflectance was considered to be Red, as shown in equation (2) in section 2.1.1. The main reason for using the CBSI-NDVI index was that it enhances the target crop in the study area to maximum, making it suitable for specific crop classification. The final output produced was rescaled to 8-bit.

4.4. Spectral Separability Analysis

The spectral separability analysis was carried out to choose the optimum number of dates required for specific crop mapping. This method was adopted because over the period, there could be a possibility of detecting spectral overlap responses from the crops in the study area. This plays a crucial role especially during level-3 classification which was carried out in phase-2 for the potato crop. Since the potato crops were divided based on their sowing period which differed by 2-3 weeks, it was important to extract the combination of those date datasets only where all the three target crop classes were at a maximum spectral distance.

In this research, three spectral separability measures were evaluated on the dataset to prepare the training data for the models. The Euclidean distance (Misra et al., 2014) was calculated that evaluated separability using the crop class mean value which is least affected by any noise present in the data. The Euclidean distance is calculated as in Eq. (4).

$$D = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2} \quad (4)$$

Here, D is the spectral distance,
n is the number of bands,
d is the DN value of pixel 'a' at the i^{th} band,
e is the DN value of pixel 'b' at the i^{th} band.

Hence, the CBSI-NDVI outputs were taken as input for evaluating separability between the classes. For evaluating spectral separability both target/interest and non-interest classes were considered. This ensured that the target class was not only separated from other target classes in the area but also separated from other classes in the area. The separability analysis using Euclidean distance was calculated for both phase-1 and phase-2 target crops.

Apart from Euclidean distance, the spectral separability was also evaluated using Transformed Divergence (TD) and Jeffries-Matusita (JM) measure. These measures were used for separability analysis for target classes of phase-2 only. The phase-1 had very small fields in the study area which resulted in the limited number of samples for each crop type. The number of pixels per crop class was not enough to calculate the TD and JM distance measures, hence Euclidean distance was calculated for phase-1 crop classes. However, the Transformed Divergence distance method resulted in the selection of a single date image for the target crops during separability analysis. But the growth of the crop is a dynamic process and at a given time, more than one crop type could yield similar spectral signatures. It is only possible to separate the crop based on their phenological cycle which could be observed by taking a combination of two or more date spectral signatures. Therefore, Transformed Divergence method results were not considered.

Lastly, the Jeffries-Matusita distance measure was evaluated for phase - 2 crop classes. Here, covariance matrices were calculated between the signatures of the classes being evaluated. The output from both Euclidean distance and Jeffries-Matusita distance measures were considered and two separate training sets were generated for the models which were further evaluated based on the classification results produced by the models. All the separability analysis measures were evaluated using ERDAS Imagine software.

4.5. Deep Learning Based Classification

In this research, two deep learning models were generated and explored for specific crop classification which was the CNN model and an integrated CNN-LSTM model. Both the models were generated using 'Keras' which is open-source software that acts as an interface to the TensorFlow python library. The model architecture used was 'sequential'. The 1D-CNN model was generated as it uses less computation and can work with limited hardware resources and a limited dataset. The model was tested against different numbers of neurons, the number of epochs, learning rates, dropout rate and activation functions. Also, the model was tested against varying training sample sizes which were around 20, 30, 40, and so on. This was done to find the minimum number that were sufficient for classification of a specific crop class and to study the performance of model for corresponding number of samples. The model architecture was optimised based on the performance of the model and classification results obtained. The optimum number of training samples and hyper-parameters were decided by interpreting the model accuracy versus the number of epochs and model loss versus the number of epochs graphs. From the graphs when training loss became approximately equal to the validation loss, it gave the optimum number of training samples. The hyper-parameters which gave the best fitting graphs were considered to be optimal for both models. The training dataset was tested against a simpler and slightly complex 1D-CNN model. The optimized model that gave the best output was then considered to be the best.

In the CNN-LSTM model, the LSTM layer was added with every 1D-CNN layer and the model was optimized using the training data. The output generated by both the classifiers gives membership value as each pixel value which is further analysed during accuracy assessment.

4.6. Fuzzy based classifier for handling heterogeneity

In an area, sowing of a particular crop by farmers depends on the conditions like availability of empty fields, seeds, or the variety of crops being sown in the field. Due to these reasons, sometimes, a particular crop type is sown by the farmers on different dates which could result in different spectral signatures from the same crop in an area. Apart from this, when a farmer owns a large field, there could be a case where a crop is sown in the whole field in consecutive days. All these reasons contribute to the difference in the growth of the same crop type introducing heterogeneity within a crop class or intra-class variation (Khitrov et al., 2020). This could further lead to the misclassification of the crop type, and thereby degrading the classification result.

The deep learning model is designed in a way that it handles the heterogeneity within a crop class by learning from the training data. So, by applying multiple convolutions over the training data and taking it to a higher dimension and by learning from the data, CNN and its hybrid models result in better classification. These models take a little more time for training. In the fuzzy domain, there are classifiers which can extract a single class at a time, in minimum time. However, fuzzy classifiers evaluates statistical parameters during the classification such as mean. Therefore, the Modified Possibilistic *c*-Means (MPCM) classifier, which is a modified version of the Possibilistic *c*-Means (PCM) classifier was evaluated and was tested against the outputs generated from deep learning model for heterogeneity within a crop class. The MPCM classifier overcomes a major limitation of the PCM classifier by generating non-overlapping clusters which were not possible using PCM classifier.

In a given input data $X = \{x_1, x_2, x_3, \ldots, x_n\}$ with 'c' no. of information classes having values $1<c<N$, the MPCM classifier tends to minimize its objective function given by eq (5)–

$$J_{MPCM}(U, V) = \sum_{i=1}^{c}\sum_{j=1}^{N} \mu_{ji} D_{ji}^2 + \eta_i \sum_{j=1}^{N} (\mu_{ji} \log \mu_{ji} - \mu_{ji}) \tag{5}$$

Here,

U = N × c matrix where c is the number of information classes (c ≥ 2) and N is the number of pixels;

V = collection set of vectors of information class centers;

Membership value, $0 \leq \mu_{ji} \leq 1$;

D_{ji} =Distance between feature vector x_k containing pixel value at k and vector v_i containing cluster centers and is computed as shown in eq (6) –

$$D_{ji} = d_{ji}^2 = ||x_j - v_i||^2 \tag{6}$$

η_i = distribution parameter computed as shown in eq (7) –

$$\eta_i = \left(\sum_{j=1}^{N} \mu_{ji\,(fcm)}^{m} D^2_{ji}\right) / \left(\sum_{j=1}^{N} \mu_{ji\,(fcm)}^{m}\right) \tag{7}$$

Here,

m = weighing exponent; for specific crop classification since one class is considered at a time, m=1;

μji for fcm is calculated as shown in eq (8) –

$$\mu_{ji\,(fcm)} = \left[\sum_{k=1}^{c} \left(\frac{D(x_j, \bar{\nu}_i)}{D(x_j, \bar{\nu}_k)}\right)^{\frac{1}{1-m}}\right]^{-1} \tag{8}$$

But for specific crop classification, m=1. Therefore, in MPCM classifier, μji is calculated as in equation (9)–

$$\mu_{ji} = \exp\left(-\frac{D_{ji}^2}{\eta_i}\right) \tag{9}$$

Here, the mean is calculated given by eq (10) in an unsupervised case or when all pixels are mixed in a given image-

$$\bar{\nu_i} = \frac{\sum_{j=1}^{N} (\bar{\mu_{ji}})^m x_j}{\sum_{j=1}^{N} (\bar{\mu}_{ji})^m} \tag{10}$$

Where x_j is the feature vector denoting the spectral response of a pixel j.

For a single class classification, the weighing exponent, i.e., m becomes equal to unity as only one target class is to be classified in the MPCM. Hence, the final objective function of MPCM classifier, as mentioned in eq (5), becomes independent of 'm'.

The training samples fed to the MPCM classifier are generated from the training fields which were identified from field data collected during field visit. In the MPCM algorithm, as a conventional approach the statistical parameter as 'mean' was calculated for the training samples. In this situation, the significance of individual sample diminishes as it's transferred and becomes highly dependent on the 'mean' of the samples. The presence of sample values close to mean decreases the significance of the outliers. Therefore, in the proposed approach for handling heterogeneity within a class, instead of calculating the mean from the input training samples, each individual samples was considered as the mean in the MPCM classifier.

In this research, output was evaluated using variance parameter to understand how much heterogeneity effect was minimized in classified output and compared for both the cases – firstly when mean of all the training sample was considered and secondly, when all the individual training samples are considered as mean. These outputs were also compared with the deep learning model outputs.

The statistical parameters such as 'mean' is one of the commonly generated parameter using the training sample data. This parameter, however, does not represent the whole range of the sample data but holds the value that lies closer to one of the sample values. During classification, a lot of sample points that belong to the same class lie far away from the 'mean' value. Therefore, these data points do not get assigned the label of that class. Therefore, a parameter such as the mean is not able to handle the heterogeneity within a class. Heterogeneity represents when there is variability in the dataset or within the class. It indicates that sample data values are distinct and are opposite of homogeneity. A homogeneous data has uniformity while heterogeneous data has non-uniform in itself. Instead of generating the mean parameter from sample data, each sample value can be treated as mean, to be used in a statistical MPCM classifier as shown in figure 10. This approach can handle heterogeneity within a class while considering the impact of each sample on classification outputs.

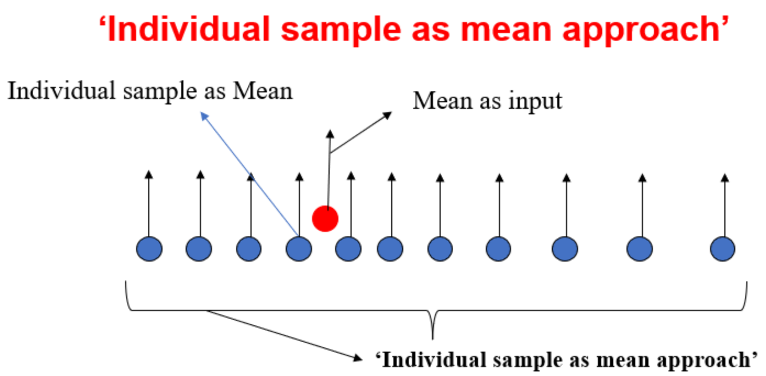


Figure 10: Concept of 'mean' and 'individual sample as mean'

4.7. Dual – Sensor Approach

In this research, the optical dataset was used as training data for the deep learning models. In one of the objectives, the microwave SAR dataset from Sentinel -1 satellite was incorporated with the optical Sentinel - 2 dataset. This dual-sensor approach was tested for phase-2 crop classes only. One of the optical date images was replaced from the training data by SAR image to generate a new training dataset. These datasets were then fed to deep learning CNN model to generate outputs. The output of the model was evaluated using MMD.

4.8. Accuracy Assessment

The performance of the deep learning models was analysed by evaluating precision, recall and f1-score from the output generated from the models (Foody et al., 1992). These parameters were evaluated from the confusion matrix generated from the classified image. The precision is calculated as shown in eq. (11) below-

$$\text{Precision} = TP/(TP+FP) \tag{11}$$

Here TP represents true positives and FP represents false negatives. True positives are the pixels that belong to the target class in the reference image and are also classified as the target class in the output (classified) image. While false positives are the pixels that belong to the non-interest (non-target) class in the reference image but are classified under the target class in the output image. The recall parameter is calculated as shown in the eq. (12) below-

$$\text{Recall} = TP/(TP+FN) \tag{12}$$

Here TP represents true positives and FN represents false negatives. False negatives are the pixels that belong to the target class in the reference image but are classified under the non-interest class in the output image. The f1-score is a function of precision and recall and can be evaluated as shown in eq. (13) below-

$$F1 = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{13}$$

The value of the f1-score lies between zero and 1. The f1-score attains a high value only when the precision and recall values are high depicting a good performance of the classifier.

The training fields in the study area were small in size and limited in number due to which there were fewer pixels identified per crop class when observed through satellite imagery. Due to the limited testing data available as a result of small fields, the number of pixels considered for evaluation of the confusion matrix was less. Therefore alternatively, accuracy assessment was carried out by evaluating Mean Membership Difference (MMD) (Singh & Kumar, 2020).

MMD is a direct and sensor-independent approach. As seen in figure 11, the target crop fields and the other crop fields were identified in the classified image produced by the models. The target crop field that was used for training the model is considered as the reference for the calculation of the MMD values. As the name suggests, the mean of difference in the membership values is calculated between the training field of the target crop and the non-training (testing) field of the target crop which gives MMD within a class. Taking Congalton's findings as a base, 75-100 pixels were selected per class (Congalton, 1991) and their average value was evaluated which was then utilized for the calculation of the MMD parameter.

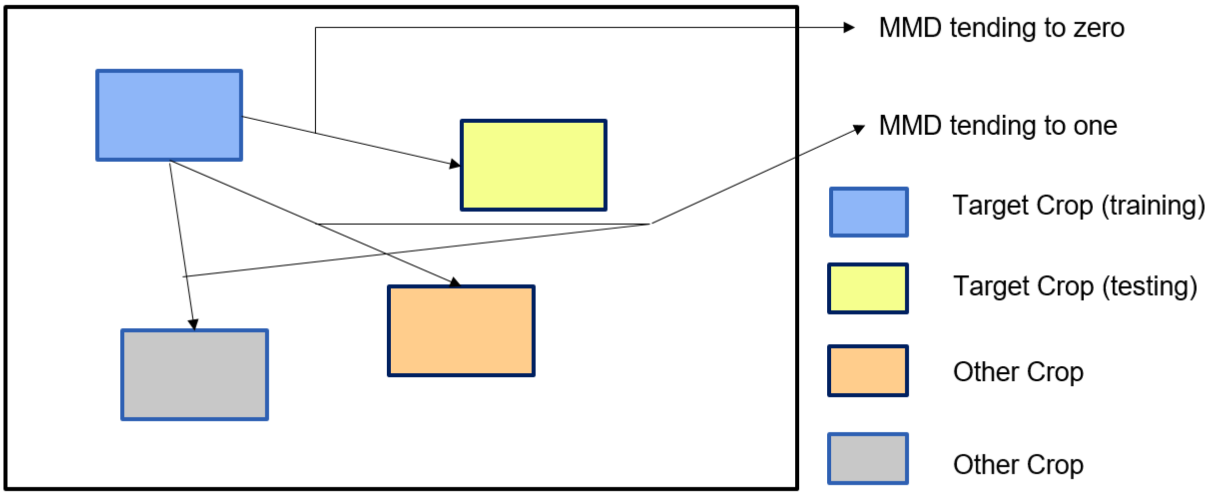


Figure 11: Calculating Mean Membership Difference

Similarly, the mean of difference in the membership values is calculated between the training field of the target crop and the fields of other non-interest classes in the area, which gives the MMD with other classes. The value of MMD lies between zero and one. If the MMD value is close to zero, it indicates that the field belongs to the target crop class and if the MMD value is close to one, it indicates the field belongs to non-interest crop class. It means, when MMD is calculated between the training and testing fields of the early potato crop, the value should be tending to zero and when the MMD is calculated between training field of early potato and the field of other (non-interest) class, the value should be tending to one. Also, the MMD approach does not require higher resolution imagery for evaluation which is essential when evaluating using fuzzy error matrix kind of assessment. The MMD value was calculated for the outputs generated from both deep learning models as well as fuzzy MPCM model and results from the dual-sensor approach.

Apart from this, assessment of the classified results for heterogeneity within a class was done by evaluating variance within a field. This parameter value indicates how homogeneously the crop fields have been classified by a classifier. The low value of variance indicates homogeneity in the output classified field and hence better classification by the classifier, and, the high value of variance indicates heterogeneity in the output classified field and hence poor classification results.

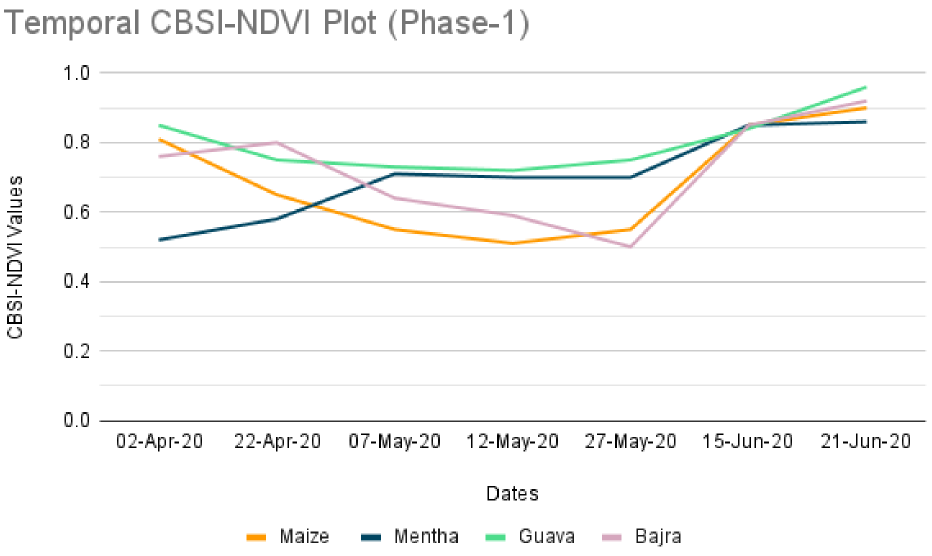
5. RESULTS AND DISCUSSION

5.1. Spectral Separability Analysis Result

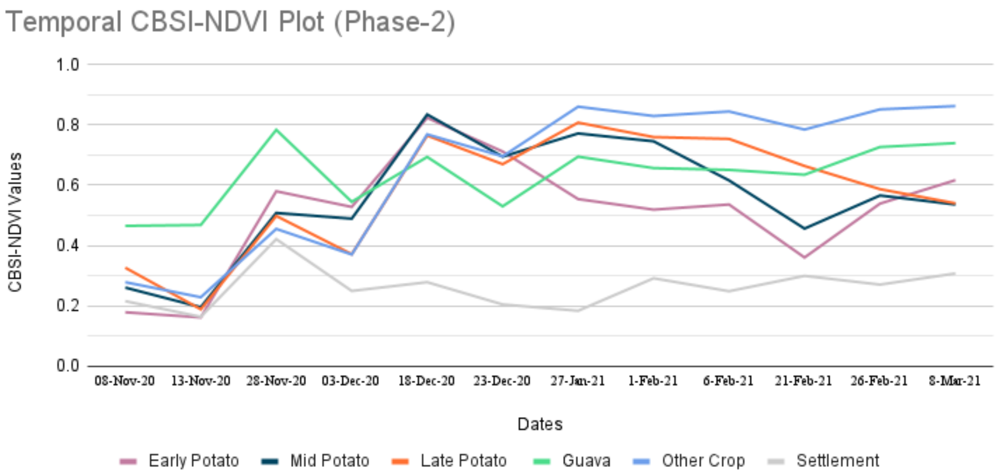
The separability analysis result gives the minimum number of dates that are suitable for specific crop mapping which addresses one of the main objectives of this research.

5.1.1. Crop Spectral Growth Curves

In this research, level-2 classification was performed in phase-1 and level-3 classification was performed in phase-2. Each crop follows a unique spectral curve over the period which plays a vital role in distinguishing the target crop from the other crops in the study area. In both the phases, each target crop has a distinctive crop cycle, meaning that all the target crops were sown on different dates resulting in distinct growth of the crop as observed in figure 12(a) and (b). The NDVI values for phase-1 crop classes were also plotted for phase-1 crop classes and shown in Appendix-I and it was observed that over time, CBSI-NDVI indices gave better separation between the crop classes. Hence, evaluation of the CBSI-NDVI indices was adopted for phase-1 and phase-2 crop classes. The details of the spectral bands considered for evaluation of CBSI-NDVI indices are mentioned in Appendix-I.



(a)



(b)

Figure 12: Temporal CBSI-NDVI plots for (a) Phase-1 crops and (b) Phase-2 crops

5.1.2. Best Temporal Date Combination

The separability analysis was performed using ERDAS Imagine software for each target/interest class and non-interest class. The separability between the interest class and non-interest class was evaluated using 3 separability measures which include Transverse divergence, Jeffries-Matusita and Euclidean distance. However, the agricultural fields for phase-1 crop classes were so small that at 10m spatial resolution of the dataset, the total number of pixels was limited in number and insufficient to evaluate spectral separability using Transverse divergence and Jeffries-Matusita measures. Therefore, Euclidean distance was used for evaluating spectral separability for phase-1 crop classes.

The CBSI-NDVI values calculated (as explained in section 2.1.1) for all the dates considered for a particular target crop were given as an input for evaluating separability. It was observed that for each target crop, the spectral distance between the target crop and non-interest crop started to saturate after a certain point. This was dependent on the number of date combinations considered to evaluate the spectral distance. A unique set of date combinations were observed for each target crop.

Tables 9 to 11 represent the spectral separability results of the target crop class with other classes (non-interest) for phase-1 in the study area evaluated using Euclidean distance. The spectral separability results evaluated using Euclidean distance between Maize (Corn) crop and the other crops were as mentioned in table 9. It was observed that after six-date combinations, the spectral separability of maize with the crop nearest to it in terms of spectral signature, which was Bajra (pearl millets), became constant. This indicated that no additional spectral information was being captured that would separate the crop class closest to maize crop, even after adding another date image into the data. The dates considered included the preparation of the field for sowing of the maize crop and the knee-high stage of the crop, which are explained in section 3.1.

Table 9: Spectral Separability analysis result of Maize (Corn) with other crop classes

No. of dates considered per combination	Dates with best min. Separation between Maize and other classes	Crop class closest to Maize during separability	Best minimum Euclidean distance measured
1 date	2	Bajra (Pearl Millets)	9
2 dates	3:7	Guava	28
3 dates	3:5:7	Bajra	35
4 dates	1:3:5:7	Bajra	39
5 dates	1:2:3:5:7	Bajra	40
6 dates	1:2:3:4:5:7	Bajra	50
7 dates	1:2:3:5:6:7	Bajra	50

From Sentinel-2, Date-1=02-04-20; Date-2=22-04-20; Date-3=07-05-20; Date-4=12-05-20; Date-5=27-05-20; Date-7=21-06-20; From Landsat-8, Date-6=15-06-20

The spectral separability results evaluated using Euclidean distance between Mentha (mint) crop and the other crops were as mentioned in table 10. It was observed that the maximum spectral separability between mentha and guava (the class closest to mentha) was given at seven-date combinations. However, while evaluating the results from the deep learning model, it was observed that six-date combinations yielded better classification results as compared to seven date combinations. Therefore, six-date combinations were selected as optimum. The optimized dates considered included the emergence of crop, vegetative growth of the crop and harvesting of the crop, which are explained in section 3.1.

Table 10: Spectral Separability analysis result of Mentha (Mint) with other crop classes

No. of dates considered per combination	Dates with best min. Separation between Mentha and other classes	Crop class closest to Mentha during separability	Best minimum Euclidean distance measured
1 date	5	Guava	22
2 dates	1:5	Guava	79
3 dates	1:3:5	Guava	82
4 dates	1:3:4:5	Guava	83
5 dates	1:3:4:5:7	Guava	93
6 dates	1:2:3:4:5:7	Guava	101
7 dates	1:2:3:4:5:7	Guava	176

From Sentinel-2, Date-1=02-04-20; Date-2=22-04-20; Date-3=07-05-20; Date-4=12-05-20; Date-5=27-05-20; Date-7=21-06-20; From Landsat-8, Date-6=15-06-20

The spectral separability results evaluated using Euclidean distance between the Guava crop and the other crops were as mentioned in table 11. It was observed that the maximum spectral separability between Guava and bajra (the class closest to guava) came out to be at seven-date combinations. There was saturation observed after four-date combinations, but the spectral signatures from the guava seem to be similar to the evergreen plants present in the study area, hence, seven-date images were considered as optimum as they provided more input in the form of spectral information. The optimized dates considered included the fruit-bearing stage of the guava trees. It covered the blooming of flowers and bearing of fruit.

Table 11: Spectral Separability analysis result of Guava with other crop classes

No. of dates considered per combination	Dates with best min. Separation between Guava and other classes	Crop class closest to Guava during separability	Best minimum Euclidean distance measured
1 date	3	Baira	21
2 dates	3:5	Baira	48
3 dates	3:4:6	Baira	94
4 dates	3:4:5:6	Baira	104
5 dates	1:3:4:5:6	Baira	104
6 dates	1:2:3:4:5:6	Baira	104
7 dates	1:2:3:4:5:6:7	Baira	106

From Sentinel-2, Date-1=02-04-20; Date-2=22-04-20; Date-3=07-05-20; Date-4=12-05-20; Date-5=27-05-20; Date-7=21-06-20; From Landsat-8, Date-6=15-06-20

Tables 12 to 14 represent the spectral separability results of the target crop class with other classes for phase-2 in the study area evaluated using Jeffries-Matusita distance. The spectral separability results evaluated using Jeffries-Matusita distance between early potato and the other crops were as mentioned in table 12. It was observed that after four-date combinations, the Jeffries-matusita distance attained its highest value possible, i.e., 1414. This indicated that all the classes considered in the study area were at the maximum spectral distance possible. Four-date combination was considered as optimum as saturation was observed in spectral distances evaluated after that. The optimized dates considered covered the tuber formation, tuber development, late harvesting of crop and field preparation for the sowing of the next crop, which is explained in section 3.1.

Table 12: Spectral Separability analysis result of early potato with other crop classes

No. of dates considered per combination	Dates with best min. Separation between early potato and other classes	Crop class closest to early potato during separability	Best minimum distance measured using Jefferies-Matusita
1 date	10	Guava	1161
2 dates	10:13	Guava	1408
3 dates	6:10:13	Mid-Potato	1413
4 dates	7:9:11:13	*	1414
5 dates	1:7:9:11:13	*	1414
6 dates	1:2:7:9:11:13	*	1414
7 dates	1:2:5:7:9:11:13	*	1414
8 dates	1:2:4:5:7:9:11:13	*	1414
9 dates	1:2:3:4:5:7:8:10:13	*	1414
10 dates	1:2:4:5:6:7:8:9:10:13	*	1414
11 dates	1:2:4:5:6:7:8:9:10:13:14	*	1414
12 dates	1:2:3:4:5:6:7:8:9:10:13:14	*	1414
13 dates	1:2:3:4:5:6:7:8:9:10:13:14:15	*	1414
14 dates	1:2:3:4:5:6:7:8:9:10:11:13:14:15	*	1414
15 dates	1:2:3:4:5:6:7:8:9:10:11:12:13:14:15	*	1414

***** All non-interest crop classes came out to be spectrally at equal distance from Early Potato

From Sentinel-2, Date-1=14-Oct-20; Date-2=24-10-20; Date-3=03-11-20; Date-4=08-11-20; Date-5=13-11-20; Date-6=28-11-20; Date-7=03-12-20; Date-8=18-12-20; Date-9=23-12-20; Date-10=27-01-21; Date-11=01-02-21; Date-12=06-02-21; Date-13=21-02-21; Date-14=26-02-21; Date-15=08-03-21

The spectral separability results evaluated using Jeffries-Matusita distance between mid-potato and the other crops were as mentioned in table 13. It was observed that after four-date combinations, the Jeffries-matusita distance attained its highest value possible, i.e., 1414. Hence, the four-date combination was considered optimum. The optimized dates considered covered the foliage, early harvesting of the crop, late harvesting of the crop and field preparation for the sowing of the next crop, which are explained in section 3.1.

Table 13: Spectral Separability analysis result of mid potato with other crop classes

No. of dates considered per combination	Dates with best min. Separation between mid-potato and other classes	Crop class closest to mid potato during separability	Best minimum distance measured using Jefferies-Matusita
1 date	10	Guava	681
2 dates	4:12	Late Potato	1339
3 dates	7:12:14	Guava	1412
4 dates	7:12:14:15	*	1414
5 dates	5:7:12:14:15	*	1414
6 dates	2:6:7:11:14:15	*	1414
7 dates	5:6:7:10:12:14:15	*	1414
8 dates	2:6:7:8:12:13:14:15	*	1414
9 dates	2:6:7:8:11:12:13:14:15	*	1414
10 dates	1:2:6:7:8:11:12:13:14:15	*	1414
11 dates	2:5:6:7:8:9:11:12:13:14:15	*	1414
12 dates	2:3:5:6:7:8:9:11:12:13:14:15	*	1414
13 dates	1:2:3:5:6:7:8:9:11:12:13:14:15	*	1414
14 dates	1:2:3:4:5:6:7:8:9:11:12:13:14:15	*	1414
15 dates	1:2:3:4:5:6:7:8:9:10:11:12:13:14:15	*	1414

***** All non-interest crop classes came out to be spectrally at equal distance from Mid Potato
From Sentinel-2, Date-1=14-Oct-20; Date-2=24-10-20; Date-3=03-11-20; Date-4=08-11-20; Date-5=13-11-20; Date-6=28-11-20; Date-7=03-12-20; Date-8=18-12-20; Date-9=23-12-20; Date-10=27-01-21; Date-11=01-02-21; Date-12=06-02-21; Date-13=21-02-21; Date-14=26-02-21; Date-15=08-03-21

The spectral separability results evaluated using Jeffries-Matusita distance between late potato and the other crops were as mentioned in table 14. It was observed that after three date combinations, the Jeffries-matusita distance attained its highest value possible, i.e., 1414 and then saturated. Hence, the three-date combination was considered optimum. The optimized dates considered covered the fallow field spectral signature being prepared for the sowing of the potato crop, and, early harvesting of the crop, which is explained in detail in section 3.1.

Table 14: Spectral Separability analysis result of late potato with other crop classes

No. of dates considered per combination	Dates with best min. Separation between late potato and other classes	Crop class closest to late potato during separability	Best minimum distance measured using Jefferies-Matusita
1 date	10	Mid Potato	986
2 dates	4:12	Other Crop	1358
3 dates	3:10:13	*	1414
4 dates	2:3:10:13	*	1414
5 dates	2:3:10:13:15	*	1414
6 dates	2:3:5:10:13:15	*	1414
7 dates	2:3:5:9:10:13:15	*	1414
8 dates	2:3:5:9:10:12:13:15	*	1414
9 dates	2:3:5:9:10:12:13:14:15	*	1414
10 dates	1:2:3:5:9:10:12:13:14:15	*	1414
11 dates	2:3:4:5:7:9:10:12:13:14:15	*	1414
12 dates	1:2:4:5:6:7:9:10:12:13:14:15	*	1414
13 dates	1:2:3:4:5:6:7:9:10:12:13:14:15	*	1414
14 dates	1:2:3:4:5:6:7:8:9:10:12:13:14:15	*	1414
15 dates	1:2:3:4:5:6:7:8:9:10:11:12:13:14:15	*	1414

***** All non-interest crop classes came out to be spectrally at equal distance from Late Potato
From Sentinel-2, Date-1=14-Oct-20; Date-2=24-10-20; Date-3=03-11-20; Date-4=08-11-20; Date-5=13-11-20; Date-6=28-11-20; Date-7=03-12-20; Date-8=18-12-20; Date-9=23-12-20; Date-10=27-01-21; Date-11=01-02-21; Date-12=06-02-21; Date-13=21-02-21; Date-14=26-02-21; Date-15=08-03-21

5.2. Deep Learning Based Classification

In this research, two deep learning models were utilized which were CNN and a hybrid of the CNN-LSTM model. This section addresses the second objective which included finding the optimized number of hyper-parameters and comparing the outputs of both deep learning models to find the model best suited for classification under conditions like having a limited dataset and smaller fields in the study area. The models were also tested against the different number of training samples to find the best training sample size as indicated in table 15-16 for CNN model and table 17-18 for CNN-LSTM model results.

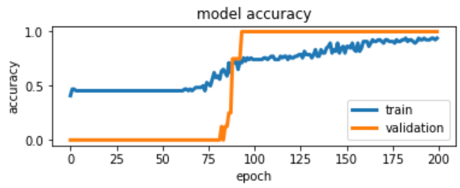
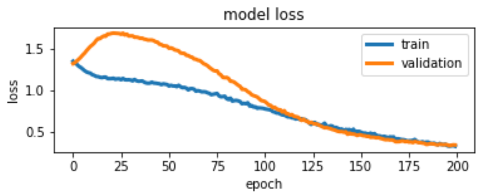
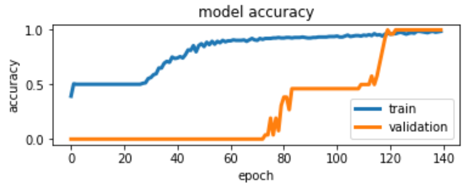
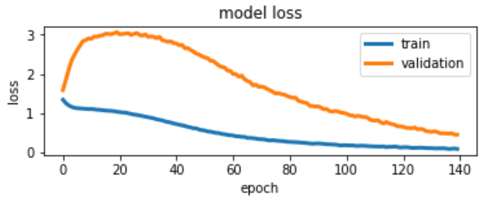
5.2.1. Optimum number of training samples

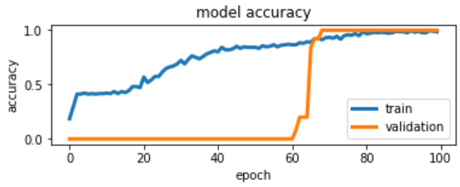
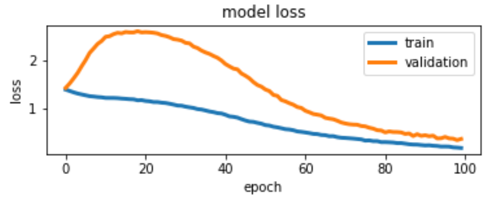
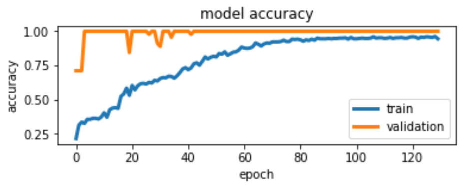
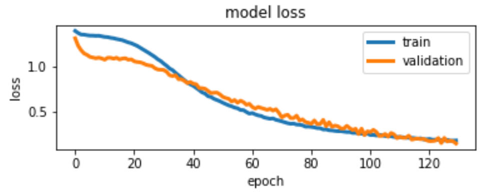
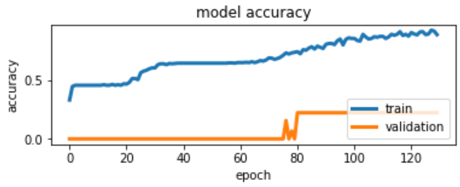
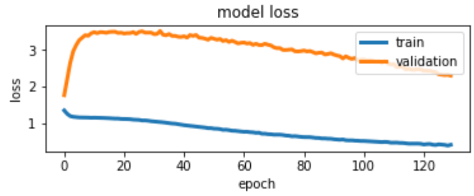
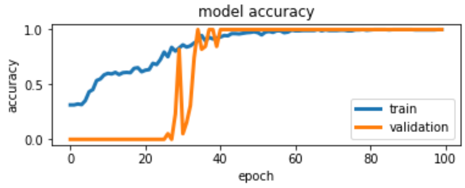
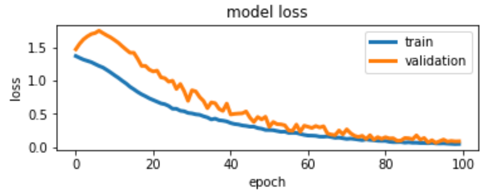
The deep learning models were tested against different number of training sample sizes. This was done to identify minimum number of training samples that were adequate in generating the best classification outputs. Also, the optimum dates selected indicated the phenological stages of the target crop that were more crucial in separating it from the non-interest classes in the study area. The optimum number of training samples suitable for each target crop was chosen by plotting the model accuracy and model loss against the number of epochs. For phase-1 target classes, the graphs were plotted for the training sample size of 30, 40, 50, 60, 70 and 100. While for phase-2 target classes, the graphs were plotted for the training sample size of 30, 40, 50 and 60.

Since the optimized dates covered different growth stages for each target crop, it was observed that the CNN and CNN-LSTM models performed better at different sample sizes. That sample size was considered to be best where the model accuracy curve and model loss curves plotted against the no. of epochs, were close to each other. When the accuracy curves of training and validation are far from each other, it indicates overfitting of the model, which is undesirable. Also, the loss curve closer to the x-axis representing the number of epochs is considered to have a good learning rate as the model loss is minimum in that case.

Table 15 shows the plots between model accuracy and model loss with no. of epochs for mentha crop while classification using the CNN model. It was observed that the best-fit curves were obtained at the sample size of 50.

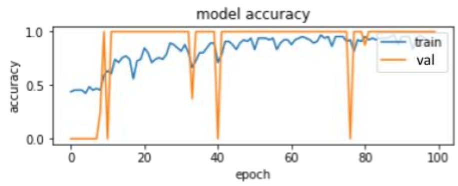
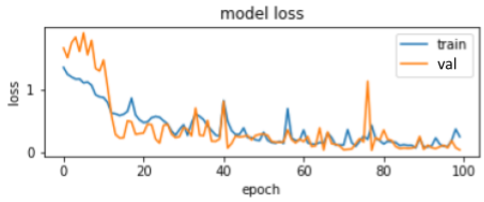
Table 15: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN model at different sample sizes for mentha.

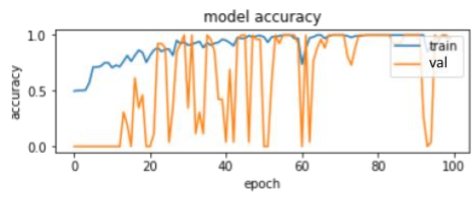
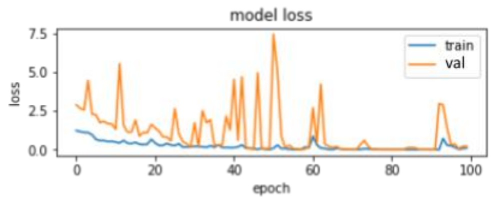
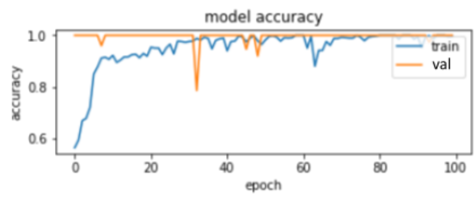
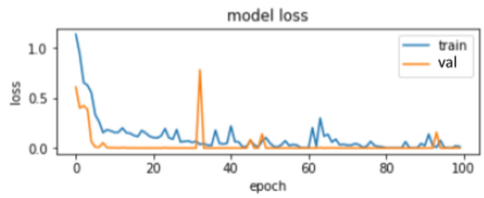
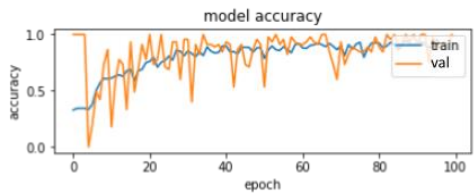
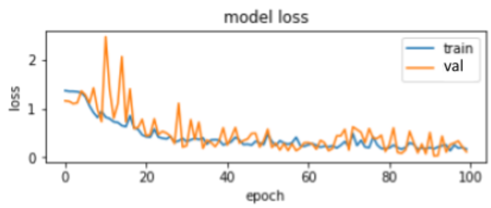
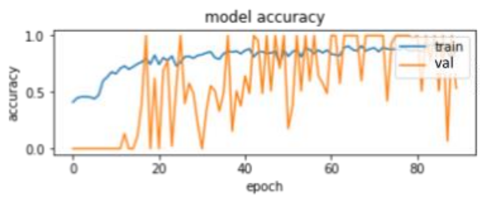
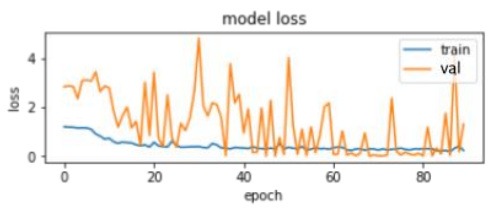
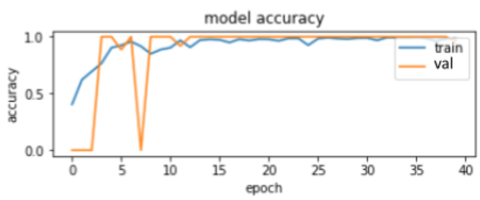
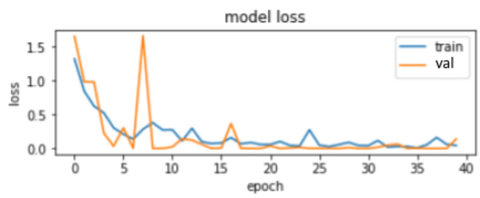
CNN Model output – Mentha		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		

50		
60		
70		
100		

The best-fit curves were plotted and observed for mentha when classified using the CNN-LSTM model. Table 16 shows the plots between model accuracy and model loss with no. of epochs for mentha crop while classification using the CNN-LSTM model. It was observed that the best-fit curves were obtained at the sample size of 60.

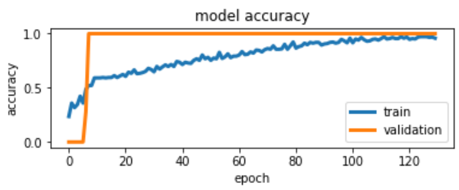
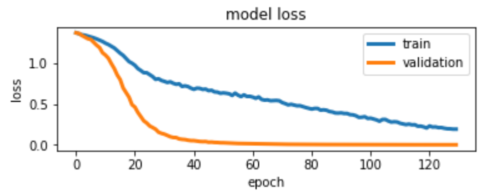
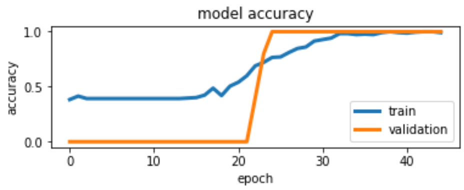
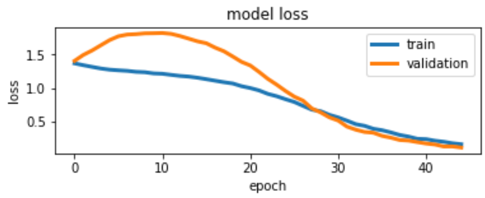
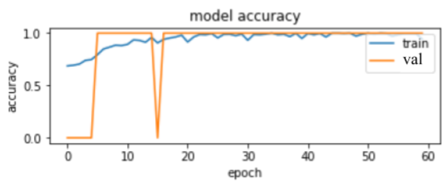
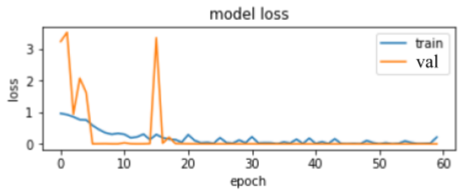
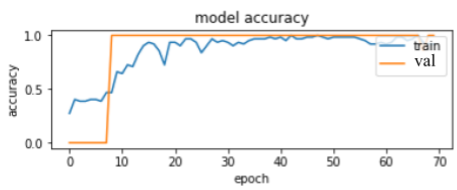
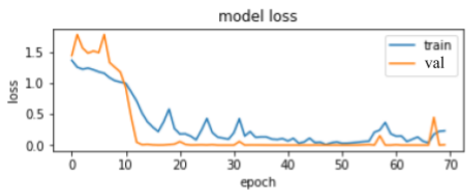
Table 16: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN-LSTM model at different sample sizes for mentha.

CNN-LSTM Model output – Mentha		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		

40		
50		
60		
70		
100		

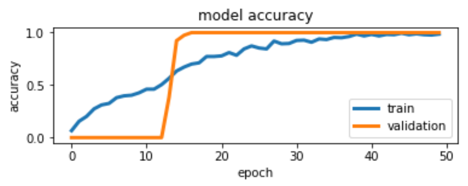
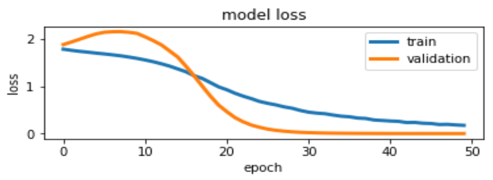
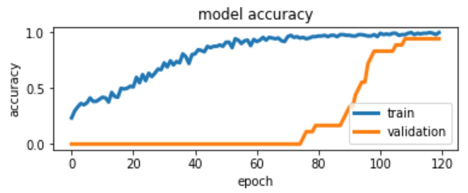
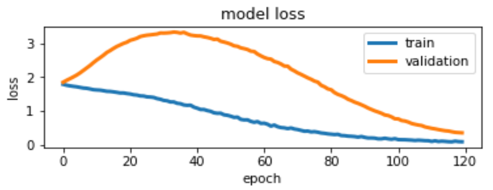
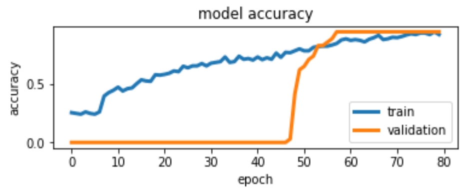
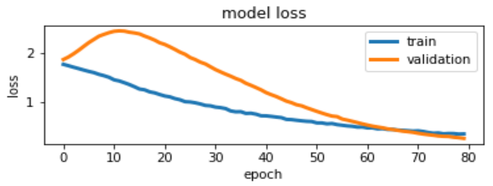
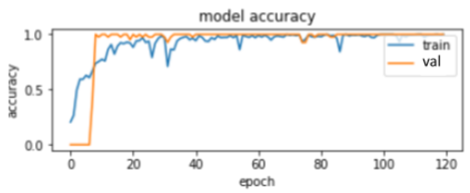
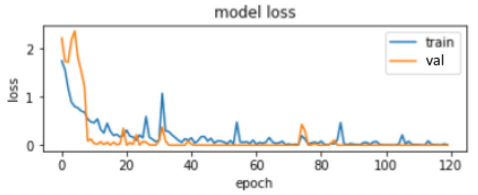
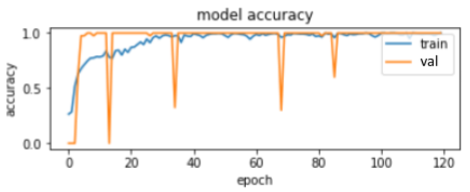
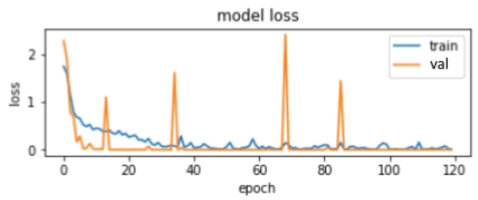
Similarly, the best-fit curves were plotted and observed for other target classes of phase-1, which are maize and guava. The best-fit curves of other phase-1 classes observed for the CNN and CNN-LSTM model along with the sample size are as shown in table 17. The other plots generated with varying numbers of training samples are shown in Appendix-II.

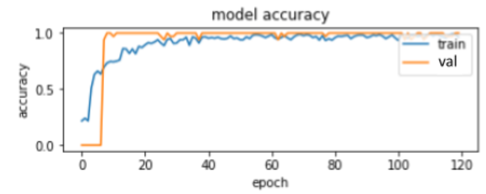
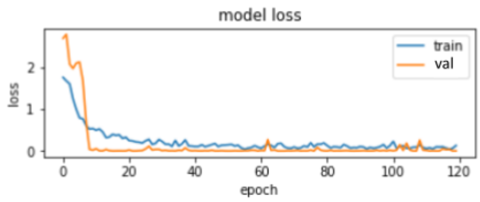
Table 17: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN and CNN-LSTM model at different sample sizes for maize and Guava.

No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
CNN Model output – Maize		
70		
CNN Model output – Guava		
40		
CNN-LSTM Model output – Maize		
100		
CNN-LSTM Model output – Guava		
30		

Similarly, the model accuracy and model loss versus no. of epochs plots were generated and plotted for all the phase-2 target classes and the sample size which produced the best-fit curves are as shown in table 18 for both CNN and CNN-LSTM models. The other plots generated with varying numbers of training samples for other target classes of phase-1 and phase-2 are shown in Appendix-II.

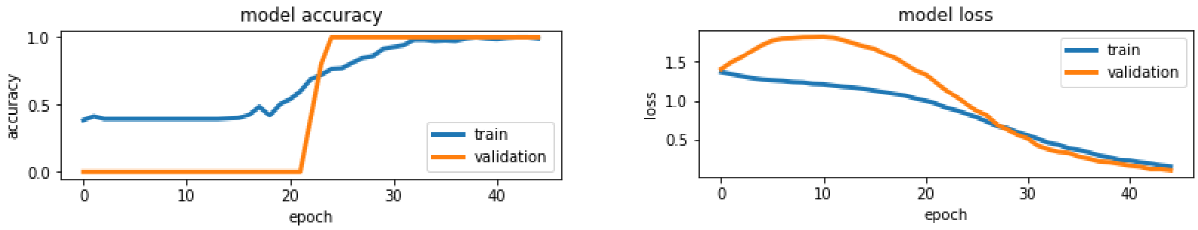
Table 18: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN and CNN-LSTM model at different sample sizes for early, mid and late potato crops.

No. of training samples	Model accuracy v/s no. of epochs	Model loss v/s no. of epochs
CNN Model output – Early Potato		
60		
CNN Model output – Mid Potato		
30		
CNN Model output – Late Potato		
40		
CNN-LSTM Model output – Early Potato		
50		
CNN-LSTM Model output – Mid Potato		
50		

CNN-LSTM Model output – Late Potato		
40		

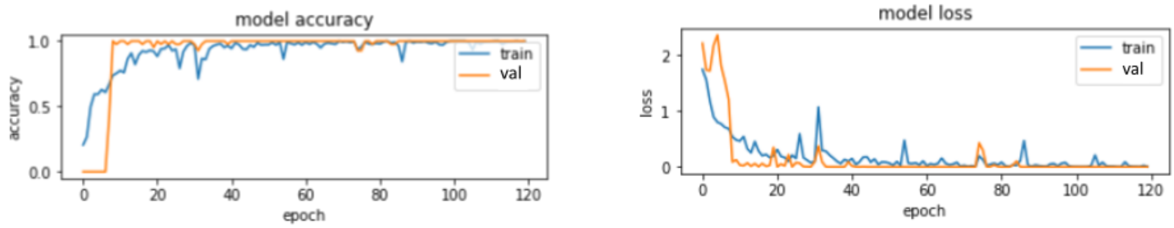
5.2.2. Optimization Results

The hyper-parameters used in the deep learning models were optimized by confirming the best-fit curves of model accuracy and model loss against no. of epochs. The deep learning models were tested along different hyper-parameters including different activation functions, no. of epochs, different batch sizes, learning rates, dropout rates, different convolutional layers and different no. of neurons. The best-fit curves obtained from CNN and CNN-LSTM models are as shown in figure 13 and 14.



(a) (b)

Figure 13: Best-fit plots for optimized CNN model between (a) model accuracy and no. of epochs (b) model loss and no. of epochs for training and validation dataset



(a) (b)

Figure 14: Best-fit plots for optimized CNN-LSTM model between (a) model accuracy and no. of epochs (b) model loss and no. of epochs for training and validation dataset

After the hyper-parameters were optimized, the structure of the deep learning model was established. The dataset was tested against two 1D-CNN models, a simple and a complex model. The simpler CNN model generated the best outputs whose structure was as shown in figure 15. It consisted of two 1-Dimensional convolution layers having 128 and 64 neurons each with Rectified Linear Unit (ReLU) activation function. It was followed by the max-pooling layer with a pool size of 1. It was followed by a dropout layer having an optimized dropout rate of 0.1. In the end, a fully connected (dense) layer was added using softmax as an activation function and Adam optimiser was used with the learning rate of 0.0001. The model was trained using a batch size of 3 while the number of epochs varied between 40 and 120 depending upon the sample size of the training data.

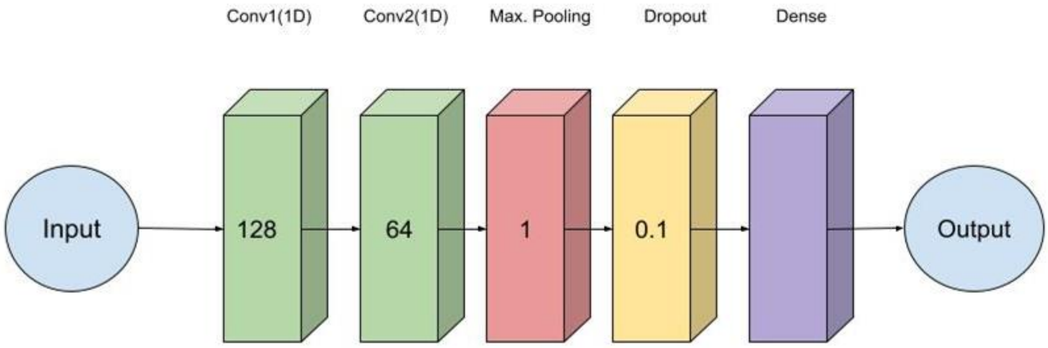


Figure 15: Optimized CNN model

In the hybrid CNN-LSTM model as shown in figure 16, an LSTM layer was included after each 1-D CNN Layer. It included a total of three CNN-LSTM layer pairs along with the ReLU activation function. The CNN layer contained 32, 64 and 128 neurons respectively while the LSTM layer contained 64, 128 and 256 neurons respectively. Each CNN-LSTM pair was followed by a dropout layer with a dropout rate of 0.6, 0.5 and 0.4 respectively. The last layer was a fully connected layer that contained a softmax activation function and Adam optimizer with a learning rate of 0.001. The model was trained with the batch size of 5 while the number of epochs varied between 70 and 140 with 120 being suitable for most of the crop classes.

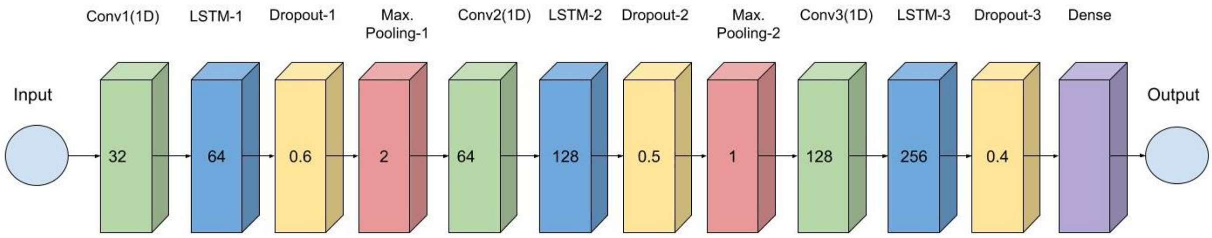


Figure 16: Optimized CNN-LSTM model

5.2.3. Accuracy Assessment through f1-score, precision and recall

The classified results generated by the CNN models were evaluated using f1-score, precision and recall and were as shown in table 19. From the results, it was observed that the f1-score of the early potato, mid potato and late potato were within the acceptable range indicating that the deep learning classifier generated fairly good outputs for the potato crops. It was observed that the precision for early potato and late potato came out to be around 0.6 indicating that most of the class labels were detected correctly by the classifier. The recall for all the three potato classes came out to be equal to one indicating that all the pixels were being classified into relevant classes. Whereas, the model does not give acceptable results based on the f1-scores observed for maize, mentha and guava crops. It was observed that there were more false positives for maize, mentha and guava resulting in lower precision which indicated that the relevant class was not being detected by the classifier during training. Also, a low value of recall was observed for maize indicating more false negatives which means that the pixels were not being classified into their relevant classes. This led to a poor f1-score for the phase-1 crop classes.

Table 19: Accuracy Assessment using precision, recall and f1-score for CNN model outputs

Target Crop Class	Precision	Recall	F1-score
Maize	0.4	0.666	0.498
Mentha	0.4	1	0.574
Guava	0.2	1	0.333
Early Potato	0.6	1	0.75
Mid Potato	0.8	1	0.88
Late Potato	0.6	1	0.75

The outputs generated from the CNN-LSTM models were evaluated using precision, recall and f1-score which were as shown in table 20. It was observed that the f1-score attained by the mid-potato crop was highest indicating better classification output. Similarly, f1-score for maize came out to be the lowest due to the low precision value. This indicated that for the maize crop, there were more false positives meaning the pixels that belonged to the non-interest class in the reference image were classified into target (maize) class in the output images.

Table 20: Accuracy Assessment using precision, recall and f1-score for CNN-LSTM model outputs

Target Crop Class	Precision	Recall	F1-score
Maize	0.2	1	0.3
Mentha	0.4	0.666	0.499
Guava	0.3	1	0.461
Early Potato	0.6	0.75	0.666
Mid Potato	0.6	1	0.714
Late Potato	0.5	0.667	0.565

Due to smaller fields and limited training data available, the number of pixels considered for evaluation of precision, recall and f1-score was less, making it less reliable for evaluating the performance of the classifier. Hence, alternate methods like the evaluation of MMD was adopted to give additional input on the performance of the deep learning models.

5.2.4. Accuracy Assessment through Mean Membership Difference (MMD)

The classification outputs from CNN and CNN-LSTM models were evaluated using MMD values. This method was adopted due to the limitation of availability of less training data as a result of small agricultural fields in the study area. This limitation resulted in a limited amount of training data restricting the proper calculation of other parameters such as f1 score, precision and recall. The evaluation of MMD values was explained in detail in section 4.8. The MMD results for phase-1 target classes for both the deep learning models were as shown in table 21.

Table 21: Mean Membership Difference for Guava, Maize and mentha crop classes

Target Crop Class (Phase-1)	CNN Model		CNN - LSTM Model	
	MMD within crop class	MMD with other crop classes	MMD within crop class	MMD with other crop class
Guava	0.0666	0.738	0.0117	0.9111
Maize (Corn)	0.25	0.6143	0.0549	0.7398
Mentha (Mint)	0.0666	0.9215	0.247	0.9215

When considering Guava as the interest (target) crop class, all the other crops were considered as the non-interest class which included maize, mentha and bajra. Here, a crop field that was not considered for training the model was considered as the interest class for calculating the MMD value. It was observed that the MMD with interest class (i.e., within crop class) came out to be tending to zero while MMD with the non-interest class (i.e., with other crop class) came out to be tending to one. In this way, MMD value was evaluated for maize and mentha crops also. However, in the case of the Maize crop, MMD with other crop classes was observed low than expected indicating slightly poor classification result CNN model which seemed to improve for the CNN-LSTM model.

Similarly, MMD results were evaluated for phase-2 target crop classes which included early potato, mid potato and late potato, and is shown in table 22. It was observed that the classification results for the early potato class came out to be most desirable for both CNN and CNN-LSTM models. This was concluded as the MMD within the class was tending to zero in the CNN model and equal to zero for the CNN-LSTM model while the MMD with other classes was tending to one for both deep learning models. However, the MMD with other classes for late potato for the CNN model output was observed to be lower than expected indicating poor classification results. Although, this parameter seemed to be improved in the CNN-LSTM model results which indicated that CNN-LSTM model resulted in better classification results as compared to the CNN model. This is also true for other crop classes as the MMD values tend to be more desirable for the hybrid CNN-LSTM model as compared to the CNN model.

Table 22: Mean Membership Difference for early potato, mid potato and late potato

Target Crop Class (Phase-2)	CNN Model		CNN - LSTM Model	
	MMD within crop class	MMD with other crop classes	MMD within crop class	MMD with other crop class
Early Potato	0.0372	0.9019	0	0.992
Mid Potato	0.1137	0.8803	0.0803	0.8215
Late Potato	0.0509	0.5362	0.0078	0.735

5.2.5. Classification Results

Apart from the evaluation using MMD values, the classification results from CNN and CNN-LSTM models were also observed through hard classification outputs. The training field used for each target crop class was observed through the hard classified outputs. Also, the training field is observed through the satellite images in FCC for better visualization and interpretation.

a) Phase-1

The hard classified outputs of CNN and CNN-LSTM models are depicted in figure 17 for maize crop class. It was observed that there were misclassifications in the CNN model results which were also reflected in the MMD results shown in section 5.2.3. However, the classification seems to be improved in the CNN-LSTM model results which were also reflected in the MMD results.

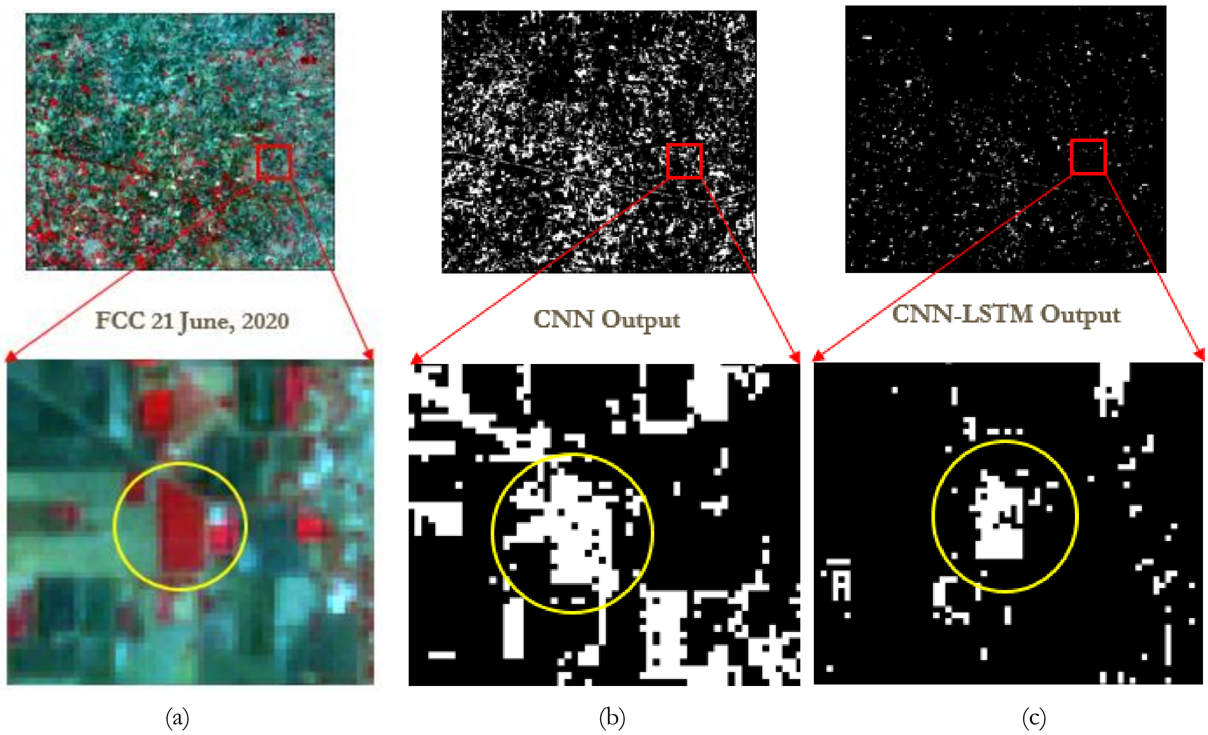
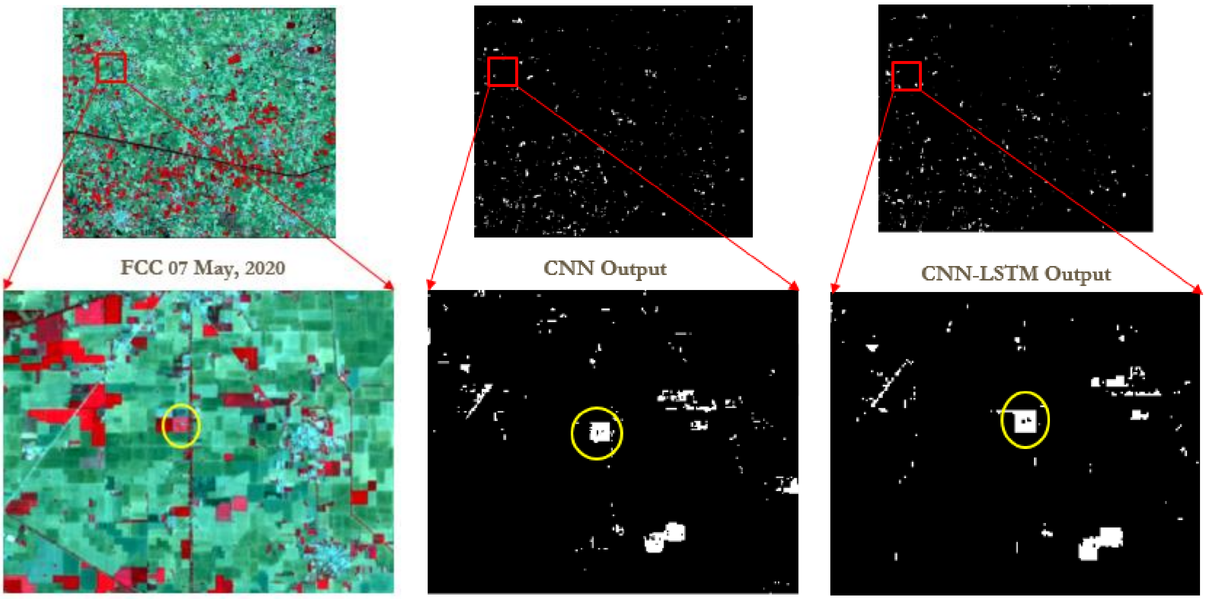


Figure 17: Maize crop in (a) FCC (b) CNN output (c) CNN-LSTM output

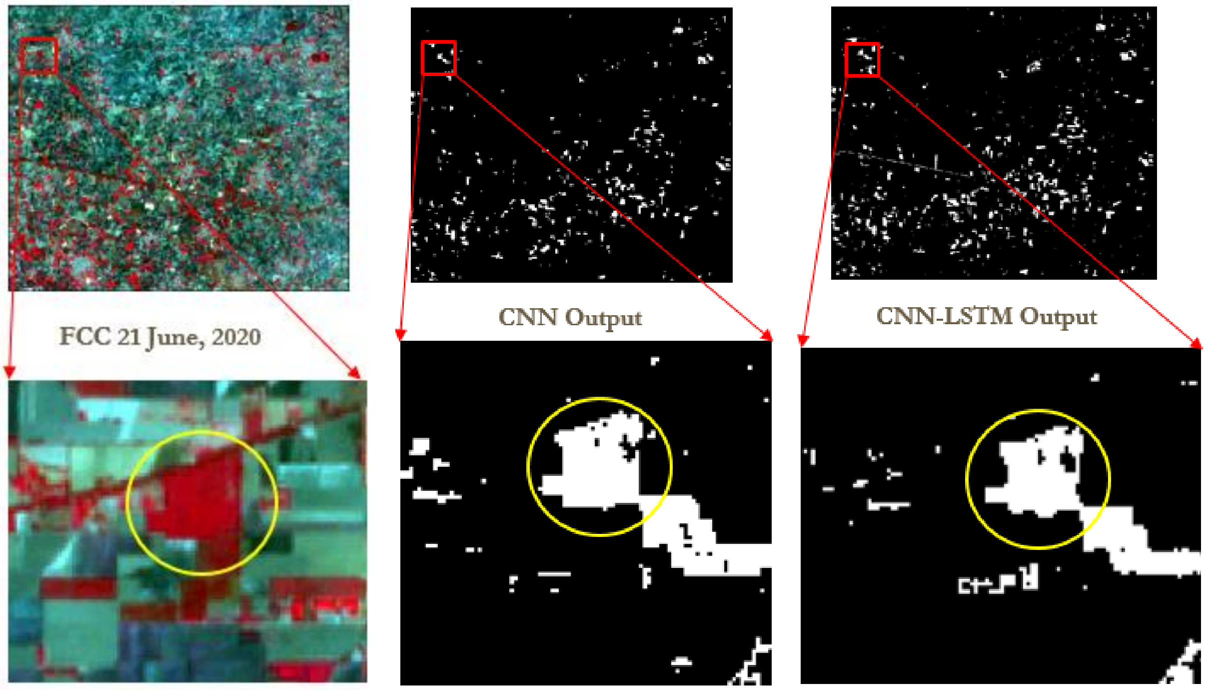
The hard classified outputs of CNN and CNN-LSTM models are depicted in figure 18 for the mentha crop class. It was observed that both the CNN model and CNN-LSTM model results indicated good classification in the study area which was also reflected in the MMD results shown in section 5.2.3. However, it could be observed that the fields classified through the hybrid model were crisp with defined boundaries indicating that the hybrid model performed better classification for mentha crop.



(a) (b) (c)

Figure 18: Mentha crop in (a) FCC (b) CNN output (c) CNN-LSTM output

The hard classified outputs of CNN and CNN-LSTM models for guava are depicted in figure 19. It was observed that both the CNN model and CNN-LSTM model results indicated acceptable classification in the study area which was also reflected in the MMD results shown in section 5.2.3. Also, it was observed that the fields classified by the hybrid model had more defined boundaries and had fewer misclassifications as compared to the CNN model output.

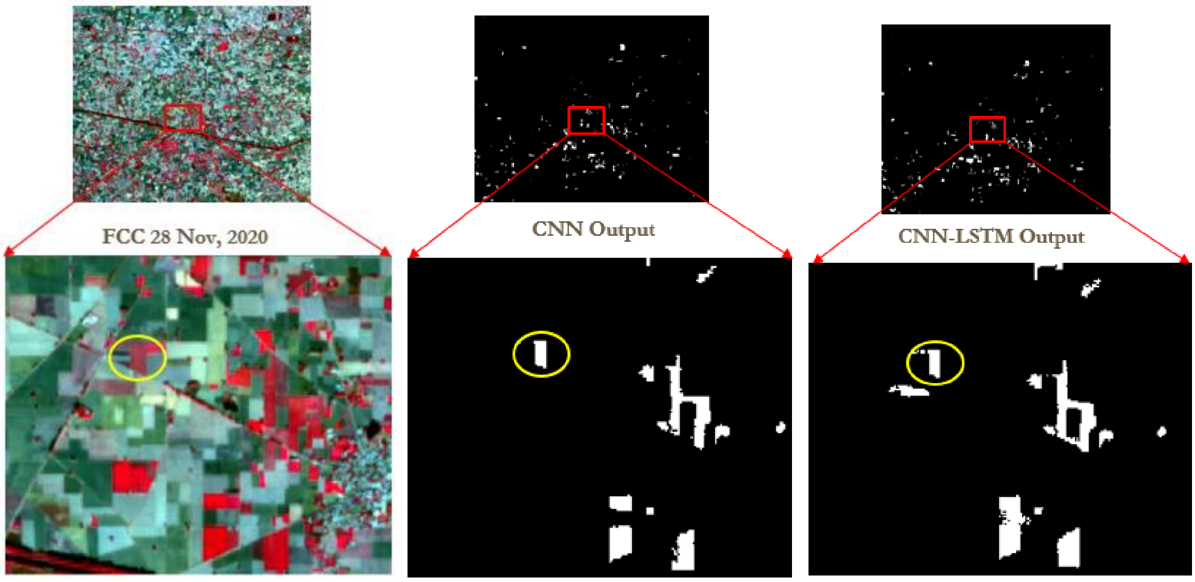


(a) (b) (c)

Figure 19: Guava crop in (a) FCC (b) CNN output (c) CNN-LSTM output

b) Phase-2

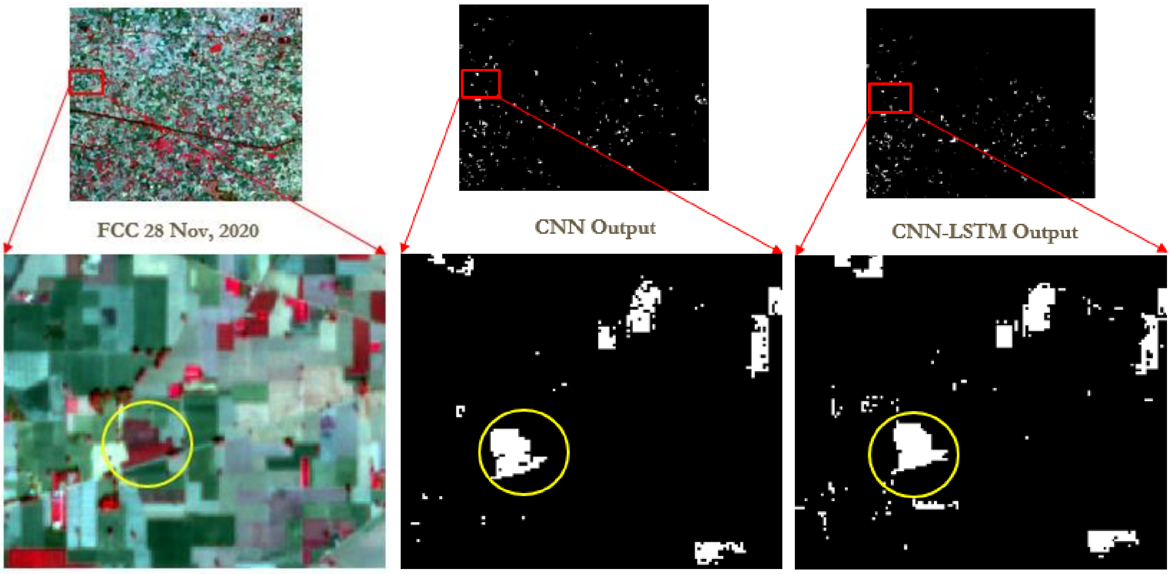
The hard classified outputs of CNN and CNN-LSTM models for early potato are depicted in figure 20. It was observed that both the CNN model and CNN-LSTM model results indicated good classification in the study area which was also reflected in the MMD results shown in section 5.2.3. Here, again the fields classified by the hybrid model were crisp and more defined as compared to the CNN model indicating better classification output.



(a) (b) (c)

Figure 20: Early potato crop in (a) FCC (b) CNN output (c) CNN-LSTM output

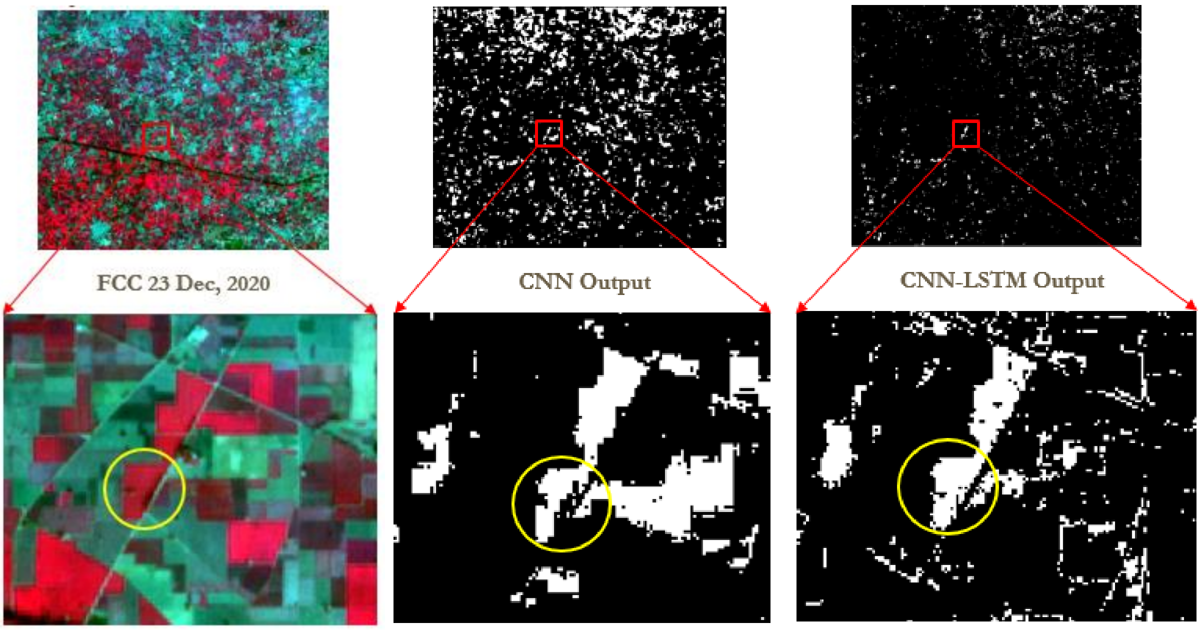
The hard classified outputs of CNN and CNN-LSTM models for mid potato class were depicted in figure 20. It was observed that both the CNN model and CNN-LSTM model results indicated good classification in the study area. From figure 21 (a), it can be noticed that the training field of the mid potato is not completely classified as there are some dark patches in the field depicting non-interest class. Whereas in figure 21 (b), it can be seen that the training field is completely classified, hence, making it a better output. Also, it is observed that the CNN-LSTM model yields some misclassified pixels when compared to the CNN model output. But when identifying the fields for a target crop, the hybrid model gives more homogeneous output making it a better classifier.



(a) (b) (c)

Figure 21: Mid potato crop in (a) FCC (b) CNN output (c) CNN-LSTM output

The hard classified outputs of CNN and CNN-LSTM models for the late potato class are depicted in figure 22. From figure 22 (a), it can be noticed that the training field of the late potato crop was not completely classified as it had some dark patches in the field. While in figure 22 (b), it can be seen that the training field is completely classified. When observed in the MMD results discussed in section 5.2.3, it indicates a low value for MMD with other classes which means that there are chances of non-interest classes to be classified into the target class, which in this case is late potato. This could be a reason why some fields are fields classified into the late potato in CNN output but not classified in the hybrid model output.



(a) (b) (c)

Figure 22: Late potato crop in (a) FCC (b) CNN output (c) CNN-LSTM output

Hence, considering the MMD results and by visually interpreting the model outputs, it could be concluded that the hybrid model yielded better classification results for both phase-1 and phase-2 crop classes, making it a better classifier.

5.3. Fuzzy based MPCM classifier

The fuzzy-based MPCM classifier was utilized in this research work, where the fuzzy classifier was made to work similar to the deep learning models as explained in section 4.6. The MPCM classifier was tested for the heterogeneity within a crop class and its outputs were compared with the outputs of deep learning classifiers. The heterogeneity within the class was observed in the training field of the target class and it was tested using the variance. The final classified output generated by the model was also tested for its accuracy using MMD values. This section 5.3 addresses the third objective of this research.

5.3.1. Accuracy assessment using variance

The MPCM model was tested against the training data for classification using two methods. As discussed in section 4.6, the fuzzy MPCM classifier evaluates using statistical parameters such as 'mean' which was further utilized to classify a particular pixel. In the first method, all the individual samples that are provided to train the MPCM model were considered as 'mean'. This implies, instead of evaluating the mean of all the input training samples, each sample was taken as a mean. By adopting this method, the MPCM classifier works like a deep learning model where each sample provided to the model was considered individually for evaluating the output. In the second method, the mean of all the input training samples was evaluated which is further utilized to classify.

To study the heterogeneity (intra-class variability), the variance was calculated within the training field in the model-generated outputs. Statistically, variance measure the spread of data. In this case, variance indicates how the membership values are spread within the training fields evaluated from the model outputs. So, a value closer to zero will indicate less spread and more homogeneity. With the increase in the variance, the classified output tends to become more heterogeneous which is undesirable. Table 23 shows the variance evaluated from the outputs of the MPCM model using the two methods explained, CNN model output and output from the CNN-LSTM model for phase-2 crop classes.

From variance results, as seen in table 23, it can be observed that the heterogeneity was best handled by the hybrid CNN-LSTM model as the values evaluated from its output were closest to zero. This means the membership values of all the pixels inside the training field of the target class lie close to each other. It was also observed that the MPCM model performed better than the CNN model when each input sample was considered as a mean while classifying the image, as could be seen in the second column of table 23. This concludes that for phase-2 crop classes, the MPCM model works similar to a deep learning model and produces a homogeneously classified field. Also, training of MPCM model is simple in comparison to CNN, CNN-LSTM models used in this research work.

Table 23: Output assessment using variance

Target Crop Class	Individual Sample as Mean in MPCM	Mean from all samples in MPCM	CNN Model Output	CNN-LSTM Model Output
Early Potato	0.2936	0.3871	7.1745	0.2645
Mid Potato	1.19	20.946	6.54	0.545
Late Potato	0.32	3.28	24.956	0.0649

5.4. Dual-sensor Approach

The dual-sensor approach was tested in this research, where one of the optical date temporal datasets was replaced by the SAR imagery to investigate if any additional information can be extracted for improving the classification. This method was tested only on phase-2 target crop classes which included early potato, mid potato and late potato. After replacing one optical image with SAR, new training data was created for each target crop which was then fed into the CNN model for classification. The evaluation of the outputs generated by the CNN was done using MMD as shown in table 24.

Table 24: Output assessment using MMD

Target Crop Class (Phase-2)	**CNN Model**	
	MMD within crop class	**MMD with other crop classes**
Early Potato	0.0254	0.25
Mid Potato	0.0941	0.211
Late Potato	0.0235	0.194

It was observed that MMD within a crop class was tending to zero indicating that the target class fields gave similar membership values. But it was observed that MMD with other (non-interest) crop classes came out to be closer to zero which is not desirable. This indicates that the other crop classes have membership values very close to that of the target class which makes it difficult to distinguish between target class and non-interest classes. Based on these results, it could be concluded that replacing an optical image with the SAR date image did not provide any additional information to the dataset but rather degraded the performance of the model. This means some of the important information supplied by the optical images got removed which resulted in poor output from the model.

6. CONCLUSION AND RECOMMENDATIONS

6.1. Conclusions

This research mainly focused on studying and designing an optimized deep learning model for specific crop classification. For this purpose, the CNN and integrated CNN-LSTM were the two deep learning models that were studied. The research was divided into two phases where distinct crop classes were studied for level-2 and level-3 classification respectively. The spatial extent of the study area was maintained in both phases. The CBSI-NDVI indices were evaluated for the satellite images which incorporated the phenological information of the crops. During specific crop classification, phenological information proved to be beneficial as it helped to reduce the spectral overlap between the crop classes in the study area.
For the first objective, the separability analysis was performed using Euclidean distance for phase-1 which extracted the minimum number of temporal images required for classification. Similarly, Jeffries-Matusita distance was used to perform separability analysis for phase-2 crop classes. The minimum number of dates selected had the maximum spectral distance between the target crop class and the nearest non-interest class in the area. This optimized the temporal domain of the dataset.
For the second objective, the CNN and integrated CNN-LSTM models were studied for the target crops of two phases. The hyper-parameters of the models were optimized to generate the best-fit model output. After evaluating the model outputs, it was observed that the integrated CNN-LSTM model produced the best classification results for the target crop classes of both phases. Also, the best classification results were yielded for phase-2 potato crop classes. This indicated that tuber formation and tuber development were important growth stages for early potato crop as it helped to separate it from other non-interest classes in the area. Similarly, foliage and harvesting stages came out to be significant for mid potato crop while tuber development and early harvesting came out to be significant for late potato crop separating them from other classes.
For the third objective, a fuzzy classifier namely Modified Possibilistic *c*-Means Classifier was used to compare the heterogeneity within a crop class in the classification results with the results of deep-learning models. After analysing the outputs from the MPCM classifier, it was observed that it performed better than the CNN model outputs but slightly poor than the integrated CNN-LSTM model. Therefore, this indicates that the MPCM classifier works similarly to a deep learning classifier while handling the heterogeneity within a crop class.
For the fourth objective, the dual-sensor approach was tested where an optical image was replaced by the SAR image in the temporal data set, to test for any additional information extracted from the dataset. This approach was only tested for phase-2 crop classes. It was observed that after replacing the image, some of the phenological information was lost as the output resulted in poor results than before. Hence for level-3 classification, the single sensor approach was better than the dual-sensor approach.

6.2. Answers to the research Questions

Question 1: What are the minimum number of temporal images required to map a specific crop while working with temporal data?
Answer: To fetch the phenological information of the crop in the study area, CBSI-NDVI spectral indices were calculated on the satellite images for each target crop. This captures the unique spectral profile of each crop in the study area observed over time. The evaluation of the separability analysis over the CBSI-NDVI indices outputs returned the maximum spectral distances between the target (interest) crop class and the non-interest crop class. The non-interest crop class indicated the class nearest to the target class in terms of spectral reflectance. It was observed that when Euclidean distance was utilized for separability of phase-1 crop classes, six dates came out to be optimum for classification of maize and mentha crops, while, for

evergreen target crops like guava, the optimum dates came out to be seven. Whereas for phase-2 crop classes, Jeffries-Matusita distance was utilized for separability which indicated four dates to be optimum for classification for early potato and mid potato classes while three dates to be optimum for late potato class.

Question 2: What are the values of the hyper-parameters that produce the best results?
Answer: For the CNN model, there were two 1-Dimensional convolution layers which consisted of 128 and 64 neurons respectively. It used Rectified Linear Unit (ReLU) as the activation function. These layers were followed by the max-pooling layer with a pool size of 1. It was followed by a dropout layer having an optimized dropout rate of 0.1. Finally, a fully connected (dense) layer was added using softmax activation function and Adam optimiser with the learning rate of 0.0001. While training the model, a batch size of 3 was used. For each target crop, the best results were achieved at varying sample sizes and hence the number of epochs of the model varied between 40 and 120.
For the hybrid CNN-LSTM model, an LSTM layer was added after each 1-D CNN layer. There was total of three CNN-LSTM layer pairs using the ReLU activation function. The CNN layer contained 32, 64 and 128 neurons respectively while the LSTM layer contained 64, 128 and 256 neurons respectively. Each CNN-LSTM pair was followed by a dropout layer with a dropout rate of 0.6, 0.5 and 0.4 respectively. Finally, a fully connected layer was added at the end containing the softmax activation function. It used an Adam optimizer with a learning rate of 0.001. The model was trained with the batch size of 5 while the number of epochs varied between 70 and 140 with 120 being suitable for most of the target crop classes.

Question 3: How much are the model loss and final accuracy for which the proposed model gives the best result for specific crop mapping?
Answer: The model loss and model accuracy were observed from the best-fit model results of the CNN and CNN-LSTM models. As depicted in figure 12, it was observed that for the CNN model the training and validation accuracy curves were very close to each other, indicating that the model has a good fitting of the dataset and hence can be applied for other crops in the different study area. Also, the loss curves of the training and validation dataset were observed to be near zero and overlapping which indicated a good learning rate. Similarly, for the best-fit CNN-LSTM model, the accuracy curves were overlapping and there was very little gap observed between the training and validation curves which indicated good fitting of the dataset. The loss curves of the training and validation dataset were observed to be near to zero and overlapping indicating a good learning rate.

Question 4: Which method best handles the heterogeneity within a crop class – deep learning model or MPCM approach?
Answer: When classifying crops, there could be cases where heterogeneity could be observed within a crop field which could be a result of sowing of crops on consecutive days, or, sowing of different varieties of crop resulting in distinct growth. To study the heterogeneity within a crop class, the variance was calculated in the field used to create the training data for the model. The outputs from CNN mode and CNN-LSTM model were compared with the outputs from two MPCM models – one where each input was considered as a mean of the samples, and the second where mean was calculated for all the input samples. It was observed that the training field in the outputs from the CNN-LSTM model came out to be most homogeneous for all the crop classes. Whereas, the MPCM model where each input was considered as the mean of the samples, proved to be better than the CNN model, making it the second-best model to handle the heterogeneity. This indicated that a statistical model such as MPCM worked and gave outputs similar to a deep learning model in terms of heterogeneity within a crop class.

Question 5: Does the SAR data give any extra information as compared to the optical dataset?

Answer: In this research, one of the optical images was replaced by the SAR dataset to test the dual-sensor approach. This approach was tested for phase-2 crop classes. The training data generated was used to produce classification outputs from the CNN model. The model outputs were then analysed by calculating MMD within crop class and MMD with other crop classes. It was observed that the value of MMD with other classes came out to be reducing which is undesirable. This means that the membership values attained by the other (non-interest) classes were very close to that of the target classes. Hence, it could be concluded that the phenological information got reduced when the SAR image was incorporated into the dataset.

Question 6: Which approach is better - single sensor or dual-sensor approach for classification?
Answer: The main aim of testing the dual-sensor approach was to find if using the SAR imagery with the optical image resulted in providing any additional phenological information about the crop which could, in turn, improve the classification of the crop. However, in the phase-2 crop classes where level-3 classification was carried out, it was observed that the dual-sensor approach didn't prove to be beneficial. The output generated by the CNN model was evaluated using MMD. It was then compared with the CNN model output from the optical dataset and observed that the classification results degraded significantly. Therefore, the single sensor approach proved to be more efficient when compared to the dual-sensor approach.

6.3. Short comings of the research

A major shortcoming of the research was the availability of less training data. Due to the presence of smaller fields in the study area and utilization of a freely available dataset of 10m resolution, the number of training pixels generated for the study was less in number. This restricted the proper evaluation of important parameters such as f1-score, precision and recall for the deep learning models. Therefore, alternate method such as calculation of MMD (Singh & Kumar, 2020) was adopted to give an additional insight into the performance of the classifier.

6.4. Recommendations and future scope

The specific crop classification is important for the yield estimation of a crop at regional, state or nation levels to meet the current food demands. This research tested two deep learning models – the 1D-CNN and CNN-LSTM model for the classification of the crops. Following were some of the recommendations for future studies:

1. The scope of classification using the deep learning models can also be tested for level-4 classification in the future based on training data available of level-4 class mapping.
2. For specific crop classification, 2D and 3D CNN models could be explored incorporating spatial information.
3. A CNN model integrated with the Gated Recurrent Unit (GRU) could be generated and tested for future classification.
4. Apart from utilizing the deep learning models for classification, the models can be further utilized for prediction purposes.

LIST OF REFERENCES

Ashourloo, D., Shahrabi, H. S., Azadbakht, M., Rad, A. M., Aghighi, H., & Radiom, S. (2020). A novel method for automatic potato mapping using time series of Sentinel-2 images. *Computers and Electronics in Agriculture*, *175*(June), 105583. https://doi.org/10.1016/j.compag.2020.105583

Assembly, U. G. (2015). *UN General Assembly, Transforming our world : the 2030 Agenda for Sustainable Development, 21 October 2015, A/RES/70/1, available at: https://www.refworld.org/docid/57b6e3e44.html [accessed 28 November 2019]*. *16301*(October), 1–35.

Assimakopoulos, J. H., Kalivas, D. P., & Kollias, V. J. (2003). A GIS-based fuzzy classification for mapping the agricultural soils for N-fertilizers use. *Science of the Total Environment*, *309*(1–3), 19–33. https://doi.org/10.1016/S0048-9697(03)00055-X

Bargiel, D. (2017). A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sensing of Environment*, *198*, 369–383. https://doi.org/10.1016/j.rse.2017.06.022

Choi, K. S., Shin, J. S., Lee, J. J., Kim, Y. S., Kim, S. B., & Kim, C. W. (2005). In vitro trans-differentiation of rat mesenchymal cells into insulin-producing cells by rat pancreatic extract. *Biochemical and Biophysical Research Communications*, *330*(4), 1299–1305. https://doi.org/10.1016/j.bbrc.2005.03.111

Chunjing, Y., Yueyao, Z., Yaxuan, Z., Liu, H., Engineering, I., District, W., … Jing, D. (2017). Application of convolutional neural network in classification of high resolution agricultural remote sensing images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII*(September), 989–992. Retrieved from https://interestingengineering.com/iot-and-smart-agriculture-are-building-our-future-cities-today

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Courville, I. G. and Y. B. and A. (1987). Deep Learning Classification of Land Cover and Crop Types using Remote Sensing Data. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-25*(3), 330–348. https://doi.org/10.1109/TGRS.1987.289804

Courville, I. G. and Y. B. and A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. In *IEEE Geoscience and Remote Sensing Letters* (Vol. 14). https://doi.org/10.1109/LGRS.2017.2681128

Foody, G. M., Campbell, N. A., Trodd, N. M., & Wood, T. F. (1992). Derivation and applications of probabilistic measures of class membership from the maximum-likelihood classification. *Photogrammetric Engineering & Remote Sensing*, *58*(9), 1335–1341.

Fukushima, K. (1988). Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, *1*(2), 119–130. https://doi.org/10.1016/0893-6080(88)90014-7

Garnot, V. S. F., Landrieu, L., Giordano, S., & Chehata, N. (2019). Time-Space Tradeoff in Deep Learning Models for Crop Classification on Satellite Multi-Spectral Image Time Series. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 6247–6250. https://doi.org/10.1109/igarss.2019.8900517

Gómez, C., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *116*, 55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Gonzalez, T. F. (2007). ImageNet Classification with Deep Convolutional Neural Networks. *Handbook of Approximation Algorithms and Metaheuristics*, 1–1432. https://doi.org/10.1201/9781420010749

Ian Goodfellow and Yoshua Bengio and Aaron Courville. (2016). *Deep Learning*. Retrieved from http://www.deeplearningbook.org

Jaskaran, S. M., & Simerjeet, K. (2020a). Package of Practices for Kharif crops. Retrieved June 23, 2021, from Punjab Agricultural University website: https://www.pau.edu/content/ccil/pf/pp_kharif.pdf

Jaskaran, S. M., & Simerjeet, K. (2020b). Package of Practices for Rabi crops. Retrieved June 23, 2021, from Punjab Agricultural University website: https://www.pau.edu/content/ccil/pf/pp_rabi.pdf

KG, Y. (2021). Potato Tuber Physiological Age, Sprouting and Emergence. Retrieved August 22, 2021, from Weekly Crop Update - University of Delaware website: https://sites.udel.edu/weeklycropupdate/?p=17943

Khitrov, N. B., Rukhovich, D. I., Koroleva, P. V., Kalinina, N. V., Trubnikov, A. V., Petukhov, D. A., & Kulyanitsa, A. L. (2020). A study of the responsiveness of crops to fertilizers by zones of stable intra-field heterogeneity based on big satellite data analysis. *Archives of Agronomy and Soil Science*, *66*(14), 1963–1975. https://doi.org/10.1080/03650340.2019.1703957

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2019). 1D Convolutional Neural Networks and Applications: A Survey. *Mechanical Systems and Signal Processing*, *151*, 1–20. Retrieved from http://arxiv.org/abs/1905.03554

Kumar, A. (2012). Liquefaction identification using class-based sensor independent approach based on single pixel classification after 2001 Bhuj, India earthquake. *Journal of Applied Remote Sensing*, *6*(1), 063531. https://doi.org/10.1117/1.jrs.6.063531

Kussul, N., Lemoine, G., Gallego, F. J., Skakun, S. V., Lavreniuk, M., & Shelestov, A. Y. (2016). Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(6), 2500–2508. https://doi.org/10.1109/JSTARS.2016.2560141

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based Learning applied to document recognition. *Proc. OF THE IEEE*, 46. Retrieved from http://ieeexplore.ieee.org/document/726791/#full-text-section

Liu, S., Qi, Z., Li, X., & Yeh, A. (2019). Integration of Convolutional Neural Networks and Object-Based Post-Classification Refinement for Land Use and Land Cover Mapping with Optical and SAR Data. *Remote Sensing*, *11*(6), 690. https://doi.org/10.3390/rs11060690

Liu, Z., & Sun, Z. (2008). Active one-class classification of remote sensing image. *International Conference on Earth Observation Data Processing and Analysis (ICEODPA)*, *7285*(December 2008), 72850L. https://doi.org/10.1117/12.816115

Loussaief, S., & Abdelkrim, A. (2018). Convolutional Neural Network hyper-parameters optimization based on Genetic Algorithms. *International Journal of Advanced Computer Science and Applications*, *9*(10), 252–266. https://doi.org/10.14569/IJACSA.2018.091031

Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). High-Resolution Aerial Image Labeling with Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(12), 7092–7103. https://doi.org/10.1109/TGRS.2017.2740362

Matton, N., Canto, G. S., Waldner, F., Valero, S., Morin, D., Inglada, J., … Defourny, P. (2015). An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series. *Remote Sensing*, *7*(10), 13208–13232. https://doi.org/10.3390/rs71013208

Mazzia, V., Khaliq, A., & Chiaberge, M. (2020). Improvement in land cover and crop classification based on temporal features learning from Sentinel-2 data using recurrent-Convolutional Neural Network (R-CNN). *Applied Sciences (Switzerland)*, *10*(1), 1–23. https://doi.org/10.3390/app10010238

Mele, B., & Altarelli, G. (1993). Dropout: A simple way to prevent Neural Networks form overfitting. *Physics Letters B*, *299*(3–4), 345–350. https://doi.org/10.1016/0370-2693(93)90272-J

Misra, G., Kumar, A., Patel, N. R., & Singh, A. (2012). Mapping specific crop- A Multi sensor temporal approach. *2012 IEEE International Geoscience and Remote Sensing Symposium*, 3034–3037. https://doi.org/https://doi.org/10.1109/igarss.2012.6350786

Misra, G., Kumar, A., Patel, N. R., & Zurita-Milla, R. (2014). Mapping a Specific Crop-A Temporal Approach for Sugarcane Ratoon. *Journal of the Indian Society of Remote Sensing*, *42*(2), 325–334. https://doi.org/10.1007/s12524-012-0252-1

Murmu, S., & Biswas, S. (2015). Application of Fuzzy Logic and Neural Network in Crop Classification: A Review. *Aquatic Procedia*, *4*(Icwrcoe), 1203–1210. https://doi.org/10.1016/j.aqpro.2015.02.153

Musande, V., Kumar, A., & Kale, K. (2012). Cotton Crop Discrimination Using Fuzzy Classification Approach. *Journal of the Indian Society of Remote Sensing*, *40*(4), 589–597. https://doi.org/10.1007/s12524-012-0201-z

Nielsen, M. (2015). *Neural Networks and Deep Learning*. https://doi.org/Sept 2015

Paola, J. D., & Schowengerdt, R. A. (1995). A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of Remote Sensing*, *16*(16), 3033–3058. https://doi.org/10.1080/01431169508954607

Paul Przyborski, R. L. (2000). *Measuring Vegetation*. Retrieved from https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php

Rawat, A., Kumar, A., Upadhyay, P., & Kumar, S. (2020). Multisensor temporal approach for transplanted paddy fields mapping using fuzzy-based classifiers. *Journal of Applied Remote Sensing*, *14*(02), 1. https://doi.org/10.1117/1.jrs.14.024524

Saritha, S., & Santhosh Kumar, G. (2015). A hierarchical framework for the classification of multispectral imagery. *Procedia Computer Science*, *46*(Icict 2014), 78–85. https://doi.org/10.1016/j.procs.2015.01.060

Sentinel-1 Instrument Payload. (2021). Retrieved August 10, 2021, from The European Space Agency website: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/instrument-payload

Sentinel missions overview. (2021). Retrieved January 25, 2021, from The European Space Agency website: https://sentinel.esa.int/web/sentinel/missions

Sharma, S., Sharma, S., & Anidhya, A. (2017). Understanding Activation Functions in Neural Networks. *International Journal of Engineering Applied Sciences and Technology*, *4*(12), 310–316. https://doi.org/10.33564/IJEAST.2020.v04i12.054

Simmons, A. S. A. B., & Chappell, S. G. (1988). Artificial Intelligence-Definition and Practice. *Journal of Oceanic Engineering*, *13*(2), 14–42. https://doi.org/0364-9059/88/04OO-0014$01 .OO

Simonneaux, V., Duchemin, B., Helson, D., Er-Raki, S., Olioso, A., & Chehbouni, A. G. (2008). The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. *International Journal of Remote Sensing*, *29*(1), 95–116. https://doi.org/10.1080/01431160701250390

Singh, A., & Kumar, A. (2020). Fuzzy-based approach to incorporate spatial constraints in possibilistic c-means algorithm for remotely sensed imagery. *International Journal of Intelligent Information and Database Systems*, *13*(2–4), 307–318. https://doi.org/10.1504/IJIIDS.2020.109459

Singh, A., Kumar, A., & Upadhyay, P. (2020). Modified possibilistic c- means with constraints (MPCM-S) approach for incorporating the local information in a remote sensing image classification. *Remote Sensing Applications: Society and Environment*, *18*(March), 100319. https://doi.org/10.1016/j.rsase.2020.100319

Stroppiana, D., Azar, R., Calò, F., Pepe, A., Imperatore, P., Boschetti, M., … Lanari, R. (2015). Integration of optical and SAR data for burned area mapping in Mediterranean regions. *Remote Sensing*, *7*(2), 1320–1345. https://doi.org/10.3390/rs70201320

Sukhmal, C., NK, P., M, A., & DD, P. (2004). Agronomy and Uses of Menthol Mint. *Proceedings of the Indian National Science Academy - Springer*, 269–297.

Sun, J., Di, L., Sun, Z., Shen, Y., & Lai, Z. (2019). County-level soybean yield prediction using deep CNN-LSTM model. *Sensors (Switzerland)*, *19*(20), 1–21. https://doi.org/10.3390/s19204363

Surinder, K. T., & Jaskaran, S. M. (2020a). Package of Practices for fruits. Retrieved June 23, 2021, from Punjab Agricultural University website: https://www.pau.edu/content/ccil/pf/pp_fruits.pdf

Surinder, K. T., & Jaskaran, S. M. (2020b). Package of Practices for Vegetables. Retrieved June 23, 2021, from Punjab Agricultural University website: pau.edu/content/ccil/pf/pp_veg.pdf

Survey, U. S. G. (2016). Landsat—Earth observation satellites. https://doi.org/https://doi.org/10.3133/fs20153081.

Survey, U. S. G. (2018). Landsat collections: U.S. Geological Survey Fact Sheet 2018. https://doi.org/https://doi.org/10.3133/fs20183049

Turkoglu, M. O., D'Aronco, S., Perich, G., Liebisch, F., Streit, C., Schindler, K., & Wegner, J. D. (2021). Crop mapping from image time series: Deep learning with multi-scale label hierarchies. *Remote Sensing of Environment*, *264*, 1–40. https://doi.org/10.1016/j.rse.2021.112603

User Guides - Sentinel-2 MSI - Overview. (2021). Retrieved August 10, 2021, from Sentinel Online, The European Space Agency website: https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/overview

Wang, S., Zhou, D., Han, X., & Yoshimura, T. (2017). Chain-NN: An energy-efficient 1D chain architecture for accelerating deep convolutional neural networks. *Proceedings of the 2017 Design, Automation and Test in Europe, DATE 2017*, 1032–1037. https://doi.org/10.23919/DATE.2017.7927142

Wu, H., & Prasad, S. (2017). Convolutional recurrent neural networks for hyperspectral data classification. *Remote Sensing*, *9*(3). https://doi.org/10.3390/rs9030298

Yang, B., & Xu, Y. (2021). Applications of deep-learning approaches in horticultural research: a review. *Horticulture Research*, *8*(1). https://doi.org/10.1038/s41438-021-00560-9

Yang, J., & Li, J. (2017). Application of deep convolution neural network. *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP 2017*, *2018-Febru*, 229–232. https://doi.org/10.1109/ICCWAMTIP.2017.8301485

Zhang, C., Pan, X., Li, H., Gardiner, A., Sargent, I., Hare, J., & Atkinson, P. M. (2018). A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *140*, 133–144. https://doi.org/10.1016/j.isprsjprs.2017.07.014

Zhang, M., Lin, H., Wang, G., Sun, H., & Fu, J. (2018). Mapping paddy rice using a Convolutional Neural

Network (CNN) with Landsat 8 datasets in the Dongting Lake Area, China. *Remote Sensing*, *10*(11). https://doi.org/10.3390/rs10111840

Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. *Remote Sensing of Environment*, *221*(November 2018), 430–443. https://doi.org/10.1016/j.rse.2018.11.032

APPENDIX-I

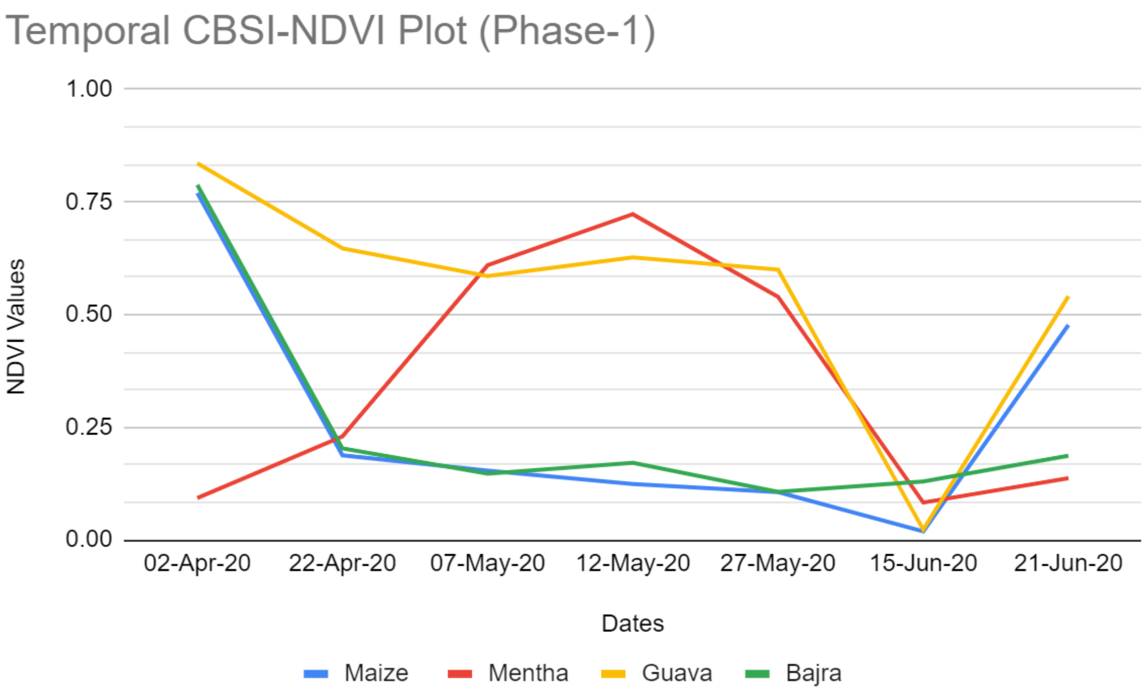


Figure A1: NDVI plot for phase-1 crop classes

Table A1: CBSI-NDVI bands considered for phase-1 crops

Class Dates	**Mentha**		**Maize**		**Guava**	
	Min.	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**
02-Apr-2020	Blue	SWIR	Blue	Narrow NIR	Blue	Narrow NIR
22-Apr-2020	Blue	SWIR	Blue	SWIR	Blue	Narrow NIR
07-May-2020	Blue	Vegetation (Red Edge)	Blue	SWIR	Blue	Narrow NIR
12-May-2020	Blue	Narrow NIR	Blue	SWIR	Blue	Narrow NIR
27-May-2020	Blue	Narrow NIR	Blue	SWIR	Blue	Narrow NIR
15-Jun-2020	TIRS 2	TIRS 1	TIRS 2	TIRS 1	TIRS 2	TIRS 1
21-Jun-2020	Blue	SWIR	Blue	Narrow NIR	Blue	Narrow NIR

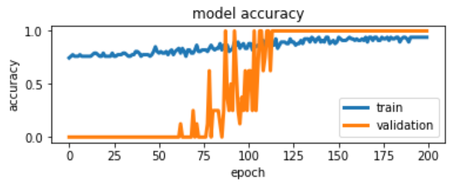
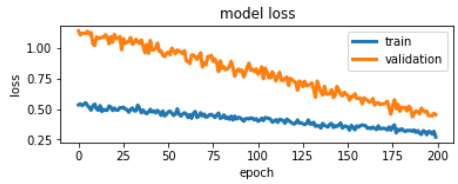
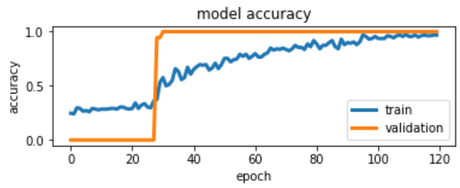
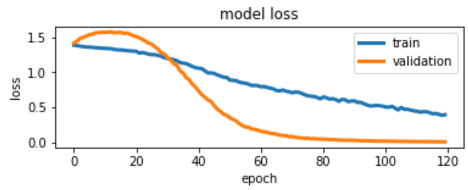
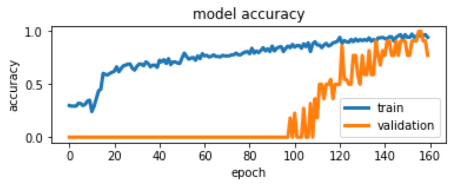
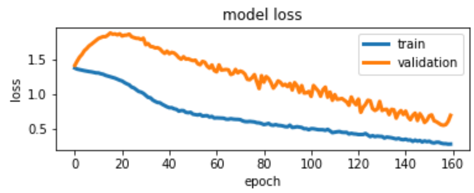
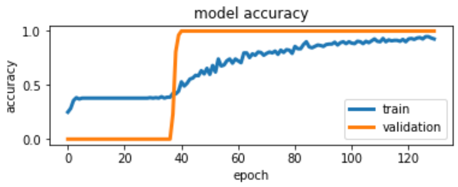
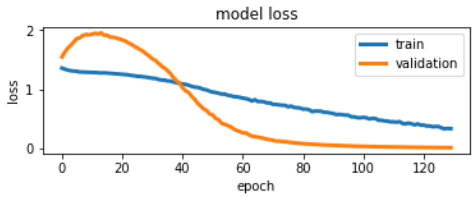
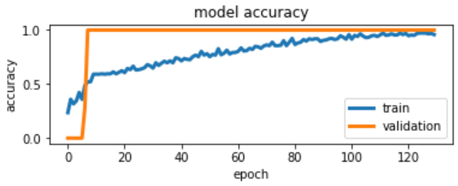
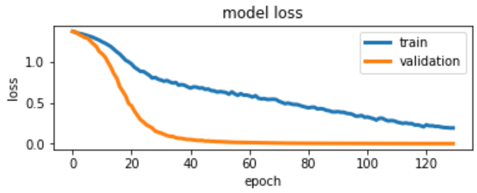
Table A2: CBSI-NDVI bands considered for phase-2 crops

Class Date	Early Potato		Mid Potato		Late Potato	
	Min	**Max.**	**Min**	**Max.**	**Min**	**Max.**
14 - Oct – 2020	Blue	Narrow NIR	Blue	Narrow NIR	Blue	Narrow NIR
24 - Oct – 2020	Blue	SWIR	Blue	SWIR	Blue	Narrow NIR
03 - Nov – 2020	Blue	SWIR	Blue	SWIR	Blue	SWIR
08 - Nov – 2020	Blue	SWIR	Blue	SWIR	Blue	SWIR
13 - Nov – 2020	Blue	SWIR	Blue	SWIR	Blue	SWIR
28 - Nov – 2020	Blue	NIR	Blue	Narrow NIR	Blue	SWIR
03 - Dec – 2020	Red	NIR	Blue	NIR	Blue	SWIR
18 - Dec – 2020	Red	NIR	Red	NIR	Red	NIR
23 - Dec – 2020	Red	NIR	Red	NIR	Red	NIR
27 - Jan – 2021	Blue	NIR	Red	NIR	Red	NIR
01 - Feb – 2021	Blue	Narrow NIR	Blue	NIR	Red	NIR
06 - Feb – 2021	Blue	SWIR	Red	NIR	Red	NIR
21 - Feb – 2021	Blue	Vegetative Red Edge	Blue	SWIR	Blue	Narrow NIR
26 - Feb – 2021	Blue	SWIR	Blue	SWIR	Blue	SWIR
08 - Mar – 2021	Blue	SWIR	Blue	SWIR	Blue	SWIR

APPENDIX-II

CNN model optimization output results for different no. of training samples.
<u>Phase-1</u>

Table A3: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN model at different sample sizes for Maize

CNN Model output - Maize		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		
50		
60		
70		

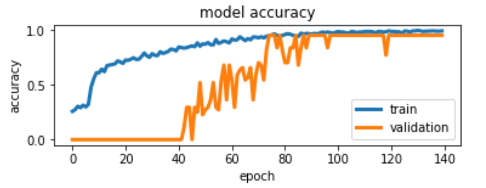
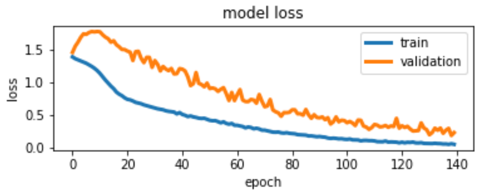
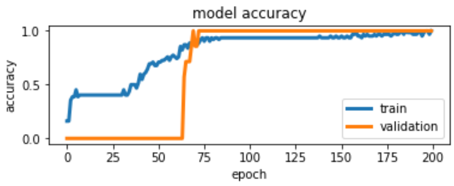
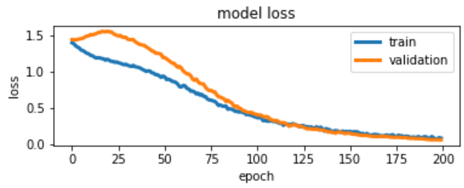
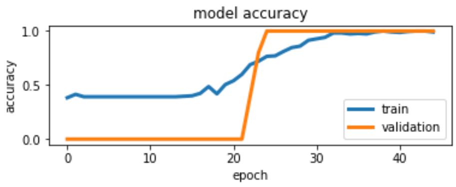
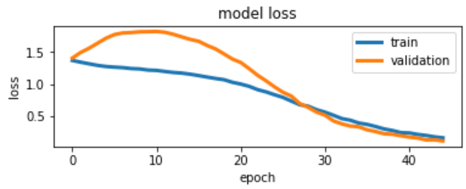
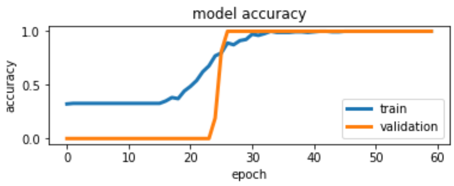
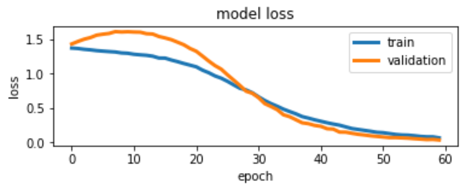
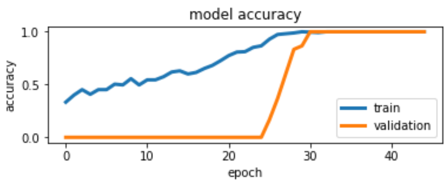
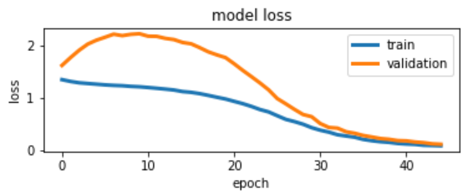
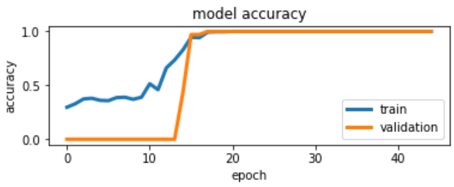
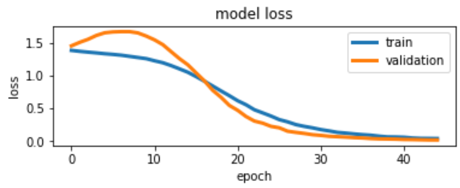
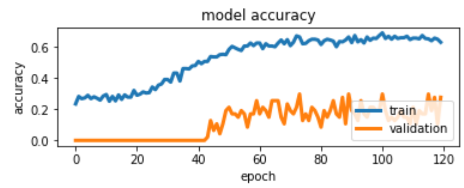
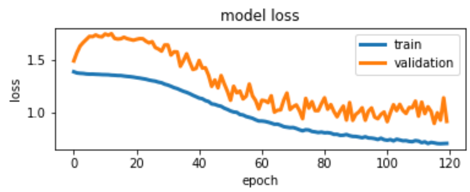
100		

Table A4: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN model at different sample sizes for Guava

CNN Model output - Guava		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		
50		
60		
70		

<table>
<tr><td>100</td><td>
</td><td>
</td></tr>
</table>

Phase-2

Table A5: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN model at different sample sizes for Early Potato

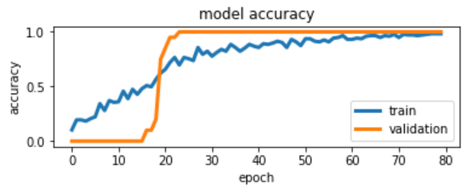
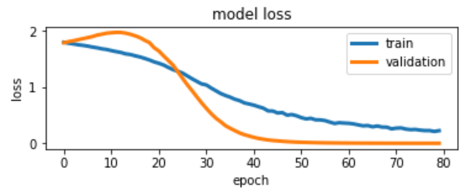
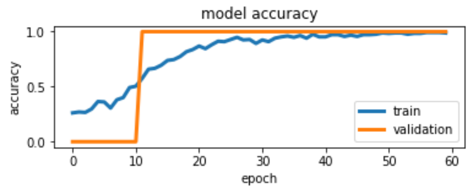
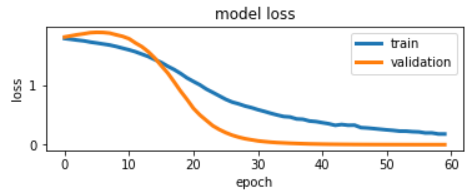
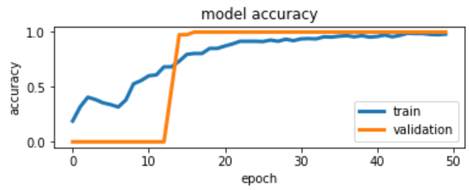
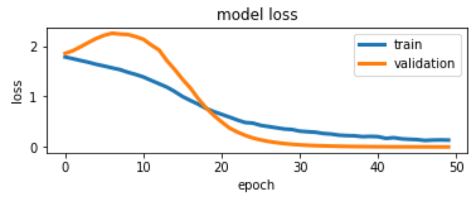
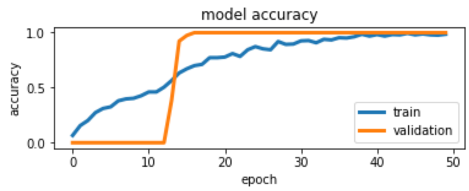
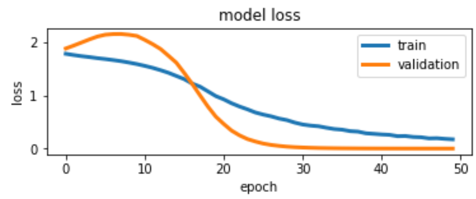
<table>
<tr><th colspan="3">CNN Model output – Early Potato</th></tr>
<tr><th>No. of training samples</th><th>Model accuracy v/s no. of epochs</th><th>Model loss v/s no. of epochs</th></tr>
<tr><td>30</td><td>
</td><td>
</td></tr>
<tr><td>40</td><td>
</td><td>
</td></tr>
<tr><td>50</td><td>
</td><td>
</td></tr>
<tr><td>60</td><td>
</td><td>
</td></tr>
</table>

Table A6: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN model at different sample sizes for Mid Potato

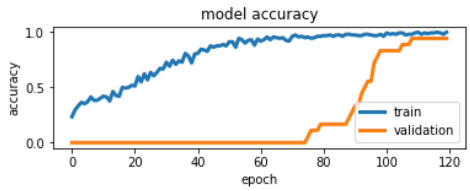
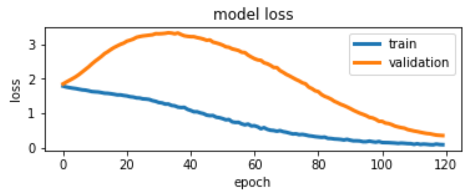
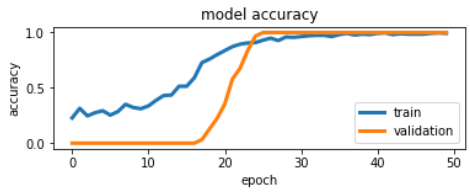
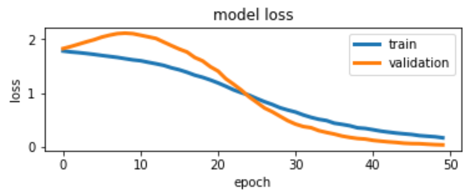
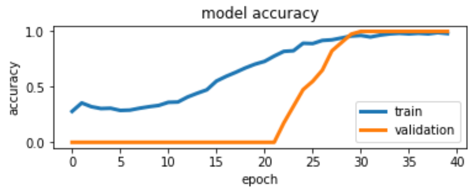
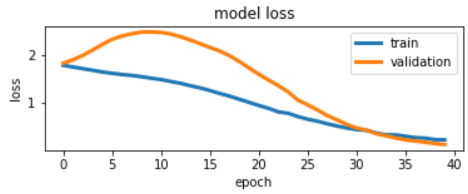
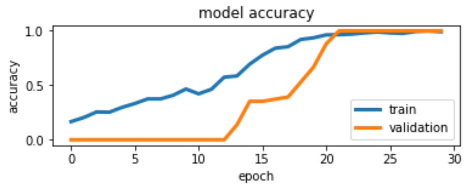
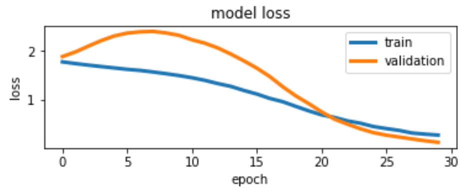
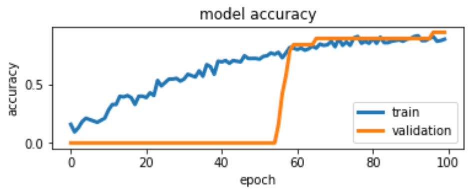
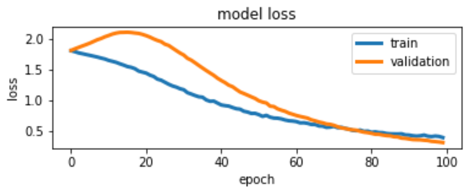
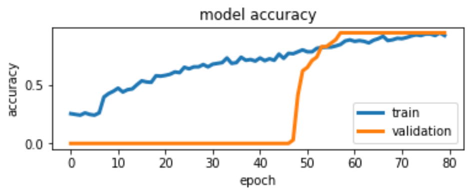
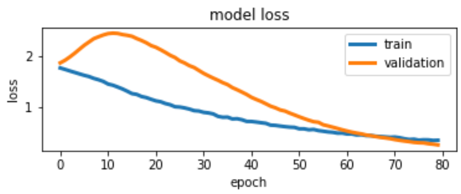
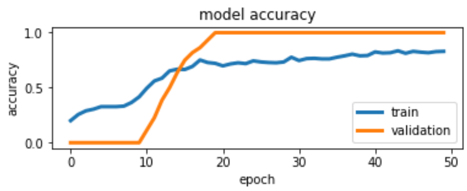
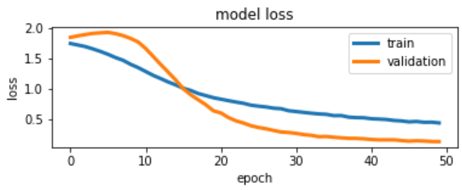
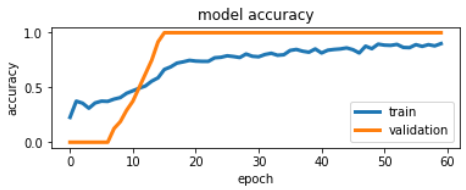
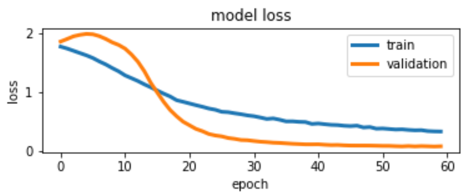
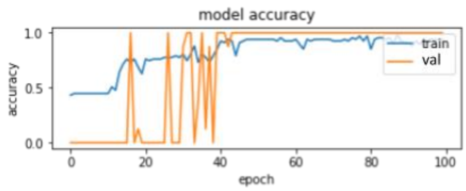
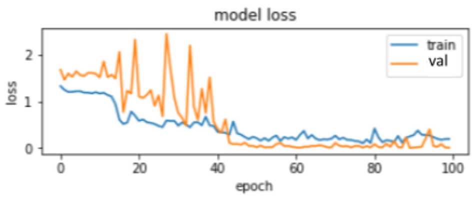
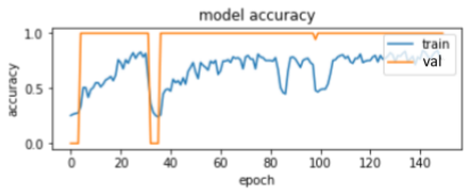
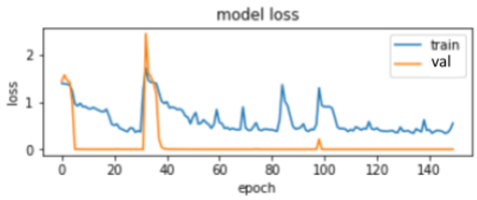
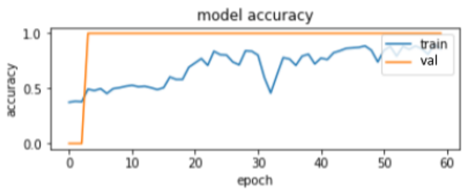
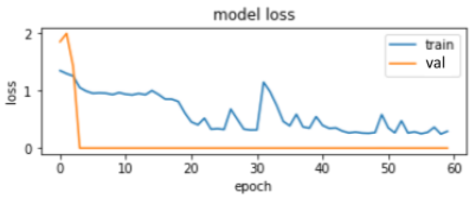
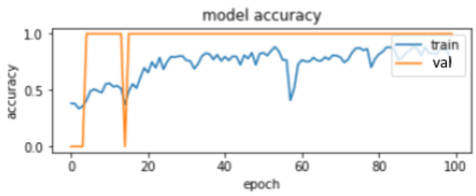
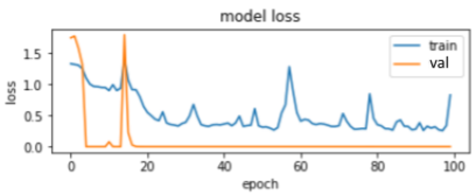
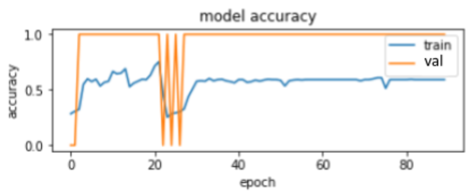
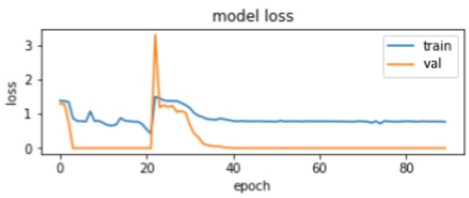
<table>
<tr><th colspan="3">CNN Model output – Mid-Potato</th></tr>
<tr><th>No. of training samples</th><th>Model accuracy v/s no. of epochs</th><th>Model loss v/s no. of epochs</th></tr>
<tr><td>30</td><td>
</td><td>
</td></tr>
<tr><td>40</td><td>
</td><td>
</td></tr>
<tr><td>50</td><td>
</td><td>
</td></tr>
<tr><td>60</td><td>
</td><td>
</td></tr>
</table>

Table A7: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN model at different sample sizes for Late Potato

<table>
<tr><th colspan="3">CNN Model output – Late Potato</th></tr>
<tr><th>No. of training samples</th><th>Model accuracy v/s no. of epochs</th><th>Model loss v/s no. of epochs</th></tr>
<tr><td>30</td><td>
</td><td>
</td></tr>
<tr><td>40</td><td>
</td><td>
</td></tr>
<tr><td>50</td><td>
</td><td>
</td></tr>
<tr><td>60</td><td>
</td><td>
</td></tr>
</table>

CNN-LSTM model optimization output results for different no. of training samples.
<u>Phase-1</u>

Table A8: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN-LSTM model at different sample sizes for Maize

CNN-LSTM Model output - Maize		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		
50		
60		
70		

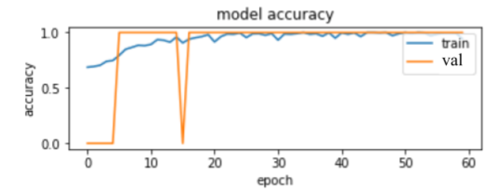
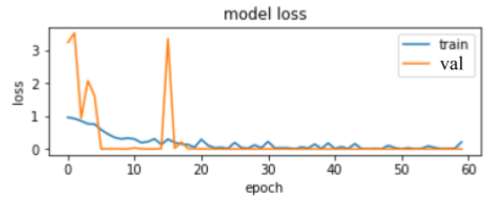
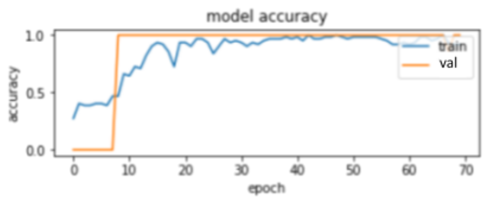
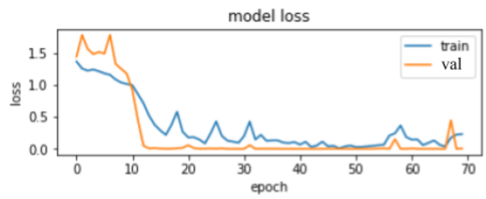
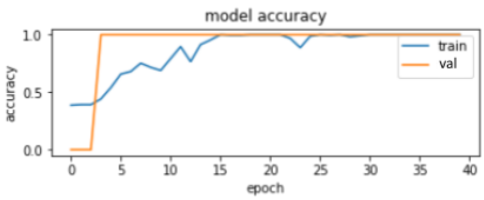
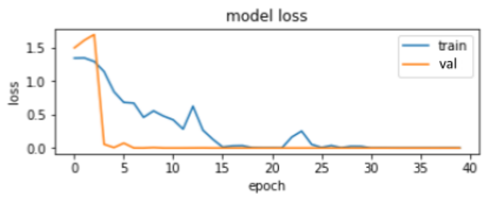
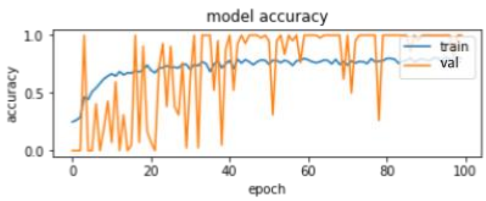
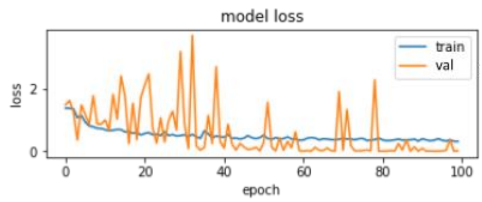
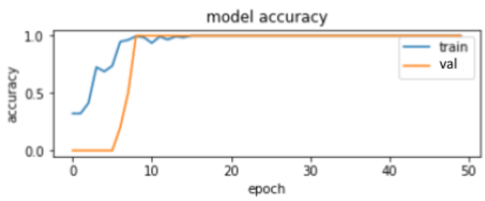
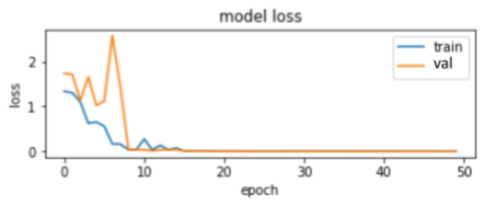
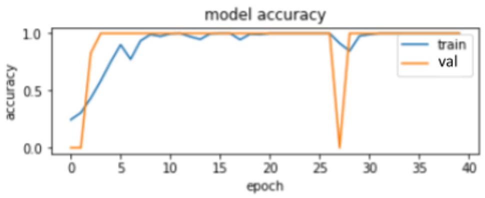
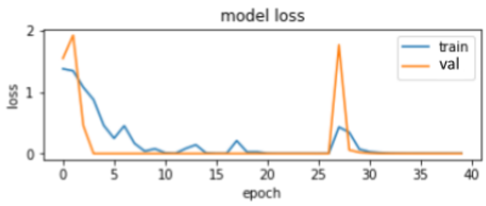
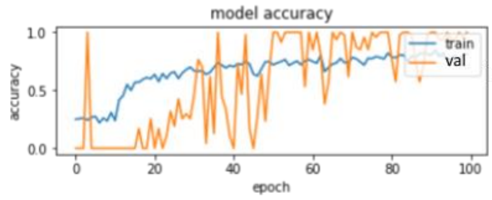
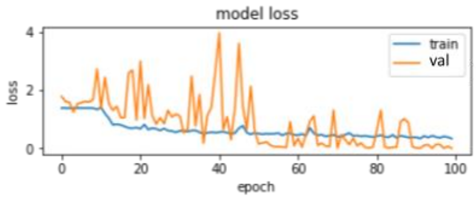
100		

Table A9: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN-LSTM model at different sample sizes for Guava

CNN-LSTM Model output - Guava		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		
50		
60		
70		

100		

Phase-2

Table A10: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN-LSTM model at different sample sizes for Early Potato

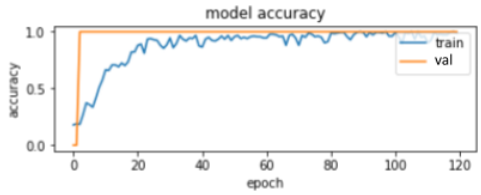
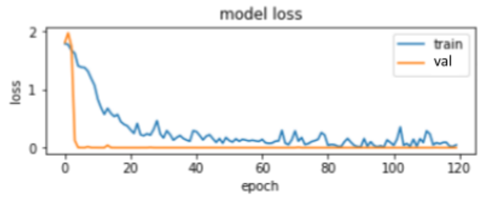
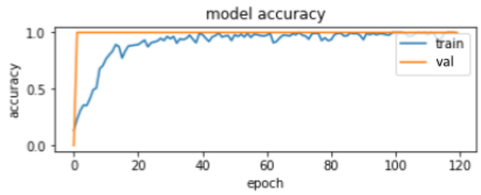
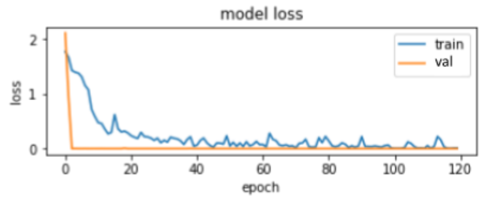
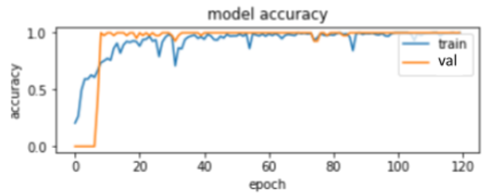
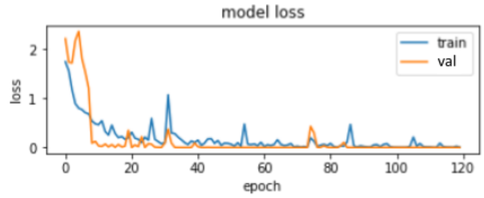
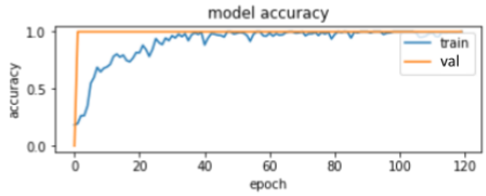
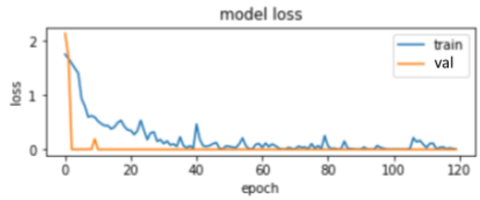
CNN Model output – Early Potato		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		
50		
60		

Table A11: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN-LSTM model at different sample sizes for Mid Potato

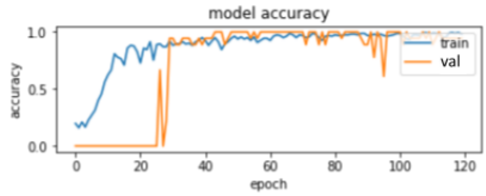
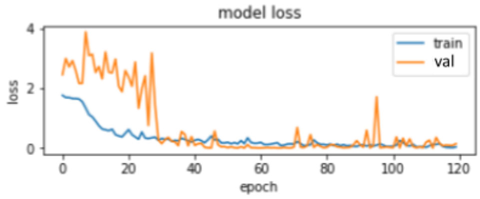
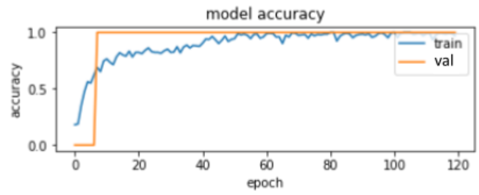
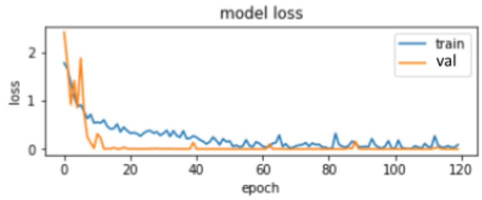
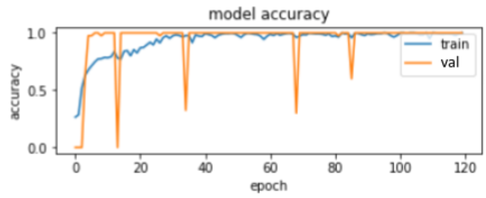
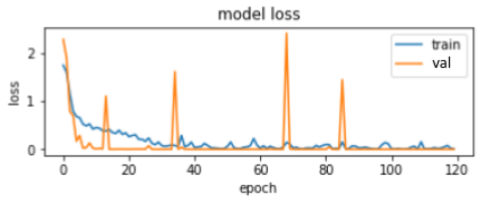
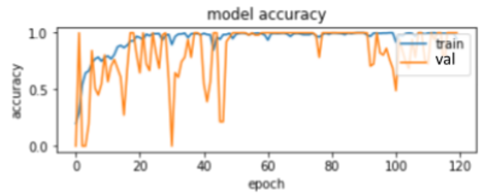
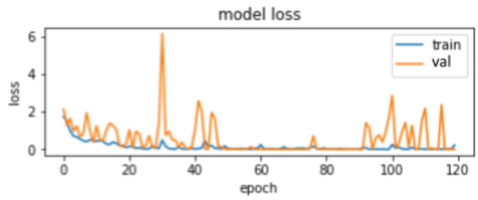
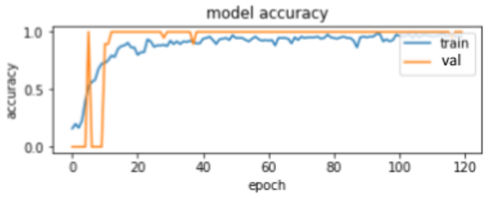
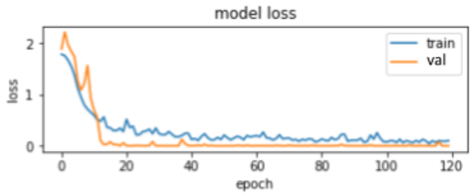
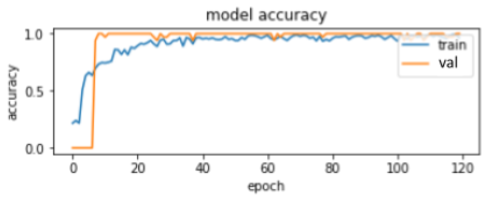
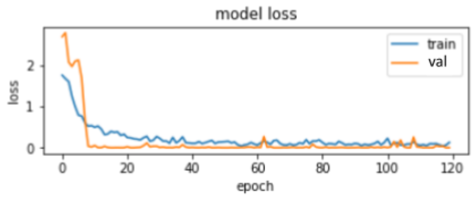
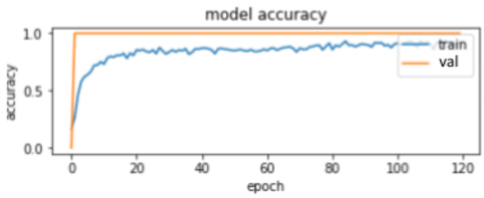
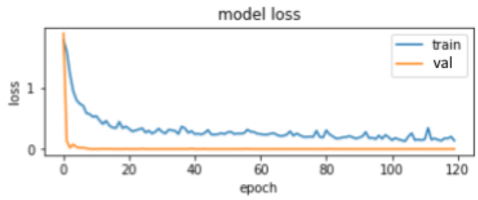
CNN Model output – Mid-Potato		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		
50		
60		

Table A12: Model accuracy versus no. of epochs and model loss versus no. of epochs plots for training and validation dataset using CNN-LSTM model at different sample sizes for Late Potato

CNN Model output – Late Potato		
No. of training samples	**Model accuracy v/s no. of epochs**	**Model loss v/s no. of epochs**
30		
40		
50		
60