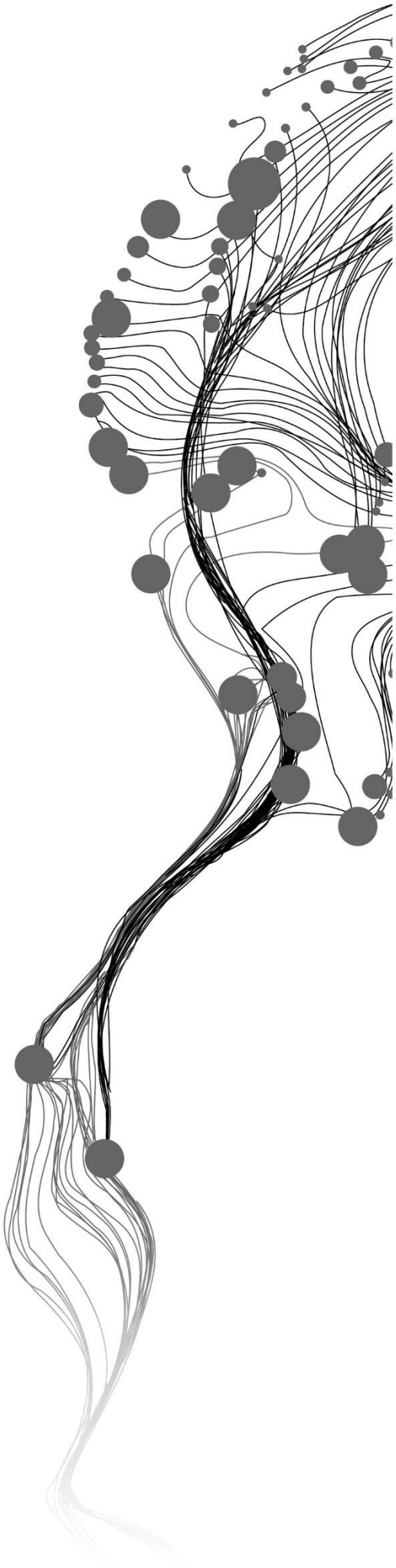
UTILIZING PHENOCAM IMAGERY AND CONVOLUTIONAL NEURAL NETWORK FOR PLANT PHENOLOGICAL STATE PREDICTION

MARIA PEREZ GALLARDO
June 2024

SUPERVISORS:
Dr, M, Khodadadzadeh
Dr, M, Farnaghi



UTILIZING PHENOCAM IMAGERY AND CONVOLUTIONAL NEURAL NETWORK FOR PLANT PHENOLOGICAL STATE PREDICTION

MARIA PEREZ GALLARDO
Enschede, The Netherlands, June 2024

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., M., Khodadadzadeh
Dr., M, Farnaghi

THESIS ASSESSMENT BOARD:
Prof. Dr., R., Zurita Milla (Chair)
Dr. rer, nat., F.J., Ellsäßer (External Examiner, ITC-NRS)
Drs. B.J., Köbben (Procedural advisor)

ABSTRACT

Phenology, the study of biological cycles in plants, is crucial for understanding ecosystem dynamics and the effects of climate change. Climate change has globally altered phenological patterns, thus impacting ecosystem interactions. While phenology research is well-established, its reliance on traditional satellite imagery presents limitations. Satellite data might have low temporal resolution, which restricts the accurate monitoring of phenological events. There is a need for reliable information using high-frequency data to gain deeper insights into these phenological events.

We propose to develop a phenological classification model to facilitate the understanding of seasonal dynamics, utilizing PhenoCam imagery known for capturing frequent and detailed observations of vegetation. Specifically, our study utilizes PhenoCam images from two stations in the University of Twente over a four-year period (2020-2023). Convolutional Neural Networks (CNNs) serve as the primary tool for image analysis, due to their ability to automatically learn and extract complex features. This approach aims to achieve accurate classification of phenological states in deciduous vegetation, focusing on spring green-up and summer green peak.

This study comprised three main phases. Firstly, we utilized PhenoCam imagery to capture comprehensive representations of phenological stages, complemented by data from other phenological sources. These images were meticulously organized and labelled to facilitate subsequent processing. Various approaches were employed to ensure accurate labelling of the images into their respective phenological stages. Secondly, a Convolutional Neural Network (CNN) model was developed to precisely identify phenological states from the labelled imagery. Lastly, we evaluated the CNN model's performance and compared it with traditional machine learning model.

The results demonstrated that the CNN model outperformed a traditional machine learning model (RF) in accurately identifying phenological states. Across all evaluated metrics, the CNN model consistently achieved an accuracy of 0.95. Overall, this study contributes to phenological research by offering a robust framework for predicting vegetation phenological stages using advanced computational techniques.

Keywords: Phenology, PhenoCam, convolutional neural network, spring green up, summer green peak, deciduous vegetation, machine learning model.

ACKNOWLEDGEMENTS

Data used in this research were provided by the PhenoCam Network particularly at the University of Twente., which has been supported by the National Science Foundation, the Long-Term Agroecosystem Research (LTAR) network which is supported by the United States Department of Agriculture (USDA), the U.S. Department of Energy, the U.S. Geological Survey, the Northeastern States Research Cooperative, and the USA National Phenology Network. We thank the PhenoCam Network collaborators, including site PIs and technicians, for publicly sharing the data that were used in this document.

Significantly, I am deeply grateful to Dr. Mahdi Khodadadzadeh and Dr. Mahdi Farnaghi for their mentorship, continuous support, and guidance throughout this thesis project. Their knowledge and insightful feedback have been crucial in helping me navigate the previously unfamiliar world of phenology and in enhancing my understanding of how we can utilize tools like deep learning. Their encouragement and expertise have significantly contributed to the successful completion of this research.

Specially, I express my profound gratitude to my parents and brothers for their unconditional support throughout this journey. From the beginning of my studies to the completion of this thesis, their constant presence, encouragement, and love, despite the geographical distance, have been invaluable. They made Panama feel close to me. Additionally, I extend my sincere appreciation to my friends both back home and from ITC for their fantastic company and joy throughout this adventure.

Without you all, this milestone would not have been possible. Thank you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

Phenology focuses on the study of the biological cycles of plants (Ge et al., 2023). It allows the detection of typical events such as blooming, changes in leaf coloration, and germination, among others(Richardson, 2023). It addresses not only individual plants but also entire ecosystems and their interactions with environmental factors. Phenology plays an important role in many sectors, such as nature conservation. Understanding how climatic factors impact phenological occurrences is essential for nature conservation, and this can be achieved through tracking phenological stages (A. D. Richardson, 2023).

Phenology has gained attention recently as an indicator of how climate change is impacting ecosystems. Phenological patterns have been altered globally by climate change, driven by factors like increased greenhouse gas emissions, deforestation, and altered land use (Ge et al., 2023). Changes in climate variables can directly impact the occurrence of phenological events, influencing ecosystem dynamics (Wang et al., 2023). Variables like temperature, precipitation, day length, solar radiation, humidity, and wind speed affect these natural processes. Shifting weather patterns have led to changes in phenological events such as flowering and leaf emergence (Lee et al., 2023).

Environmental factors play a crucial role in driving phenological changes. Extended periods of certain conditions can impact the duration of leaf activity and carbon uptake in forests (Meng et al., 2021). Changes in precipitation patterns, including more frequent and intense droughts, can decrease soil moisture levels, posing challenges for tree species adapted to specific environmental conditions. These changes may lead to declines or shifts in distribution among affected species (Schädel et al., 2023).

Phenological networks are essential for monitoring and understanding the impacts of climate change on ecosystems, biodiversity, and agriculture (Templ et al., 2018). They promote collaboration, engage citizen scientists, and provide valuable data for scientific research and conservation efforts (Templ et al., 2018). For instance, the Pan European Phenological Network (PEP725) serves as a robust source of ground truth phenological data across Europe (Templ et al., 2018).

Satellite imagery is widely used in phenological research due to its wide coverage and the ability to monitor large-scale environmental changes. However, some limitations can be addressed. One of the primary drawbacks of satellite images is the limited temporal resolution. Satellites typically pass over a specific area at regular intervals, and the revisit time can range from a few days to weeks, leading to gaps in capturing phenological events, especially those that occur rapidly or over short periods of time (Wang et al., 2023). Second, cloud cover and pollution are significant challenges in acquiring reliable satellite data, these factors can introduce discrepancies in the data and affect the quality of observations (Ge et al., 2023). Moreover, acquiring high-quality satellite imagery can be expensive, particularly for high-resolution datasets.

A PhenoCam network is an arrangement of PhenoCams located strategically across different locations to capture time-lapse images of vegetation and monitor phenological changes. These networks are designed to provide comprehensive and distributed coverage, allowing to observe and analyse plant life cycle in different environments (Richardson, 2023). PhenoCams provide high-frequency image capture, recording near-surface images every 30 minutes. They commonly capture colour images in the visible spectrum (RGB) (A. Richardson, 2023).

Capturing the dynamics of phenological events requires consistent data collection intervals. Time series data offer a chronological sequence of observations, facilitating the tracking of phenological states (Lee et al., 2023). They enable analysis of shifts in phenological patterns over extended periods, aid in detecting anomalies from typical patterns, and contribute to assessing the occurrence and rate of phenological shifts. This information is crucial for decision-making, ensuring timely responses to changing environmental conditions (Lou et al., 2023).

Phenological processes often involve complex and nonlinear patterns, influenced by various environmental factors. While statistical methods have their utility in certain aspects of phenological research, machine learning (ML) methods provide capabilities in handling complex and high-dimensional data (Guo et al., 2021). ML provides a data-driven approach to phenological studies, relying on algorithms to automatically learn and extract meaningful information from large datasets (Guo et al., 2021). Studies have demonstrated that ML methods, can outperform traditional statistical methods in various domains, estimating and monitoring flowering phenological stages is one of them (Zhu et al., 2022).

Deep learning models (DL), when well-designed and trained on relevant data, have demonstrated state-of-the-art performance in various image-related tasks. This high performance makes them a promising approach for accurate and reliable phenological classification (Capinha et al., 2021). Several characteristics of DL models make them particularly effective when dealing with large datasets, as they can capture complex patterns and features present in diverse datasets (Capinha et al., 2021).

One of these powerful tools is convolutional neural networks (CNNs), which are known for accurately classifying phenological stages captured in images. CNNs are designed to automatically learn hierarchical spatial features from images and can recognize patterns regardless of their location in the image. The convolutional layers in CNN models extract features from images, representing meaningful patterns in the input data, and these features are automatically learned during the training process(Kattenborn et al., 2021). CNNs have been applied to a broad range of thematic categories. Most of the applications correspond to agricultural studies, with a smaller scale focusing on conservation and forestry (Kattenborn et al., 2021). Several phenological studies have been conducted using CNN models. One of these studies fused MODIS and OLI images to map vegetation growth days and key phenological stages for mixed vegetation regions in China, using vegetation indices (Kun et al., 2023). This study highlights the capacity of CNNs in extracting phenological features.

1.2. Research gap

Phenological events can occur rapidly, posing challenges for their accurate monitoring using traditional satellite imagery, which operates at lower frequencies (Song et al., 2024). This limitation results in gaps that impact precise observation of these dynamic transitions. Ensuring reliable information on phenological changes is crucial for effective ecological monitoring, highlighting the need for approaches that can offer more detailed and frequent insights into these ecological shifts (Song et al., 2024).

Common approaches to monitor phenological patterns include observational networks like PEP725. However, these data have some drawbacks such as spatial distribution, location accuracy and citizen scientific experience (Tian et al., 2021). PhenoCam images offer the capability to monitor phenological changes at multiple levels, ranging from forests to individual trees, and even down to leaf and branch scales, serving diverse research purposes (Song et al., 2024). These images provide consistent and continuous data, facilitating detailed analysis across various spatial scales. Despite the presence of numerous stations aimed at monitoring deciduous vegetation, there remains a significant gap in leveraging PhenoCam technology for

a comprehensive exploration and understanding of seasonal dynamics in these ecosystems(Richardson, 2023).

The application of machine learning techniques, particularly deep learning, using PhenoCam images is an emerging area of interest in phenological research (Richardson, 2023). While existing studies have primarily focused on deep learning for phenological classification tasks using satellite images, the potential of applying these techniques to PhenoCam imagery remains underexplored (Katal et al., 2022). Deep learning has demonstrated promising results in automating the analysis of phenological stages and dynamics from satellite data (Katal et al., 2022). However, leveraging PhenoCam images presents an opportunity to enhance accuracy, which is crucial for monitoring rapid phenological changes at finer spatial scales and improving ecological insights (Richardson, 2023).

The application CNNs with PhenoCam imagery represents a promising approach in phenological research, providing valuable insights into the seasonal dynamics of vegetation. This thesis aims to contribute to the field of phenology through this application, offering a reliable framework for future studies. The outcomes are intended to serve as a benchmark in environmental management and conservation efforts.

1.3. Objective

Develop a phenological classification model for deciduous trees in the University of Twente region, applying CNN with PhenoCam images from 2020 to 2023. The model aims to classify the images into two key phenological states: spring green-up and summer green peak, enabling a comprehensive understanding of the seasonal dynamics over this period.

In order to accomplish this objective, the following subobjectives and corresponding research questions will guide this study:

1.3.1. Subobjectives

1. Data collection

    - To gather and ensure the quality of PhenoCam images spanning the four-year period (2020-2023) for the study area, along with phenological data from additional resources.

    - To assign labels to the images based on their respective phenological states.

2. CNN model development

    - To develop a CNN model that identifies phenological states for deciduous trees, specifically spring green-up and summer green peak, for a four-year period (2020-2023) by classifying images into these two key phenological states.

3. Model assessment

    - To assess the performance of the developed CNN model in identifying phenological states for deciduous trees and compare it with traditional machine learning models.

1.3.2. Research questions

1. Data collection

- How can PhenoCam images spanning the four-year period (2020-2023) be effectively gathered and organized alongside supplementary phenological data?
- Which approaches and additional datasets could be utilized to label PhenoCam images based on their phenological states?

2. CNN model development

- To what extent does the choice of architecture and parameters for the CNN model influence its ability to identify the phenological states?

3. Model assessment

- How does the developed CNN model identify the phenological states for deciduous trees?
- How does the CNN model's performance in identifying the phenological states compare with traditional machine learning models?

2. STUDY SITES AND DATASETS

2.1. Study site

The study area incorporates two PhenoCam stations situated within the campus of the University of Twente, located in Enschede, the Netherlands (Figure 1). The PhenoCam network is a collaborative network of digital cameras that continuously captures images of vegetation. It was established in 2008 and currently consists of more than 700 sites worldwide (A. Richardson, 2023).

The campus of the University of Twente is situated in an urban environment, providing an ideal representation for studying phenological patterns within an urban context. By examining deciduous trees in this setting, we gain valuable insights into vegetation dynamics and phenological processes.

The campus is 146 acres, and the region has a temperate climate with mean monthly temperatures from 2 to 18 °C, annual precipitation varies between 505 mm and 897 mm and the mean monthly relative humidity from 73% to 89%, according to KNMI station. With an elevation of 45 meters.

The study area boasts a diverse and vibrant mix of vegetation, which primary vegetation is deciduous broadleaf and secondary grassland, the dominant specie is Quercus. The campus landscape is covered with a variety of deciduous tree species (Schelhaas et al, 2014) , contributing to the rich ecological mosaic. The foliage undergoes dynamic transformations throughout the seasons, ranging from the vivid greens of spring and summer to the warm hues of autumn.

The study area is characterized by the presence of Quercus species, commonly known as oak trees. Deciduous trees shed their leaves annually and exhibit distinct seasonal changes that can be clearly observed in the spring and summer seasons (Gričar et al., 2022). These two seasons, spring and summer, are the focus of this study.

In early spring, oak trees begin the process of bud break, where dormant buds swell, and new leaves start to emerge. This period is crucial as it marks the transition from winter dormancy to active growth (Gričar et al., 2022). As the season progresses, the new leaves expand, and the tree canopy becomes denser. The fresh leaves are typically bright green, contributing significantly to the overall greenness of the landscape (Gričar et al., 2022).

By early summer, oak trees have usually developed a full canopy of mature leaves. The dense foliage provides ample shade, the leaves reach full maturity, displaying a deep green colour (Gričar et al., 2022).

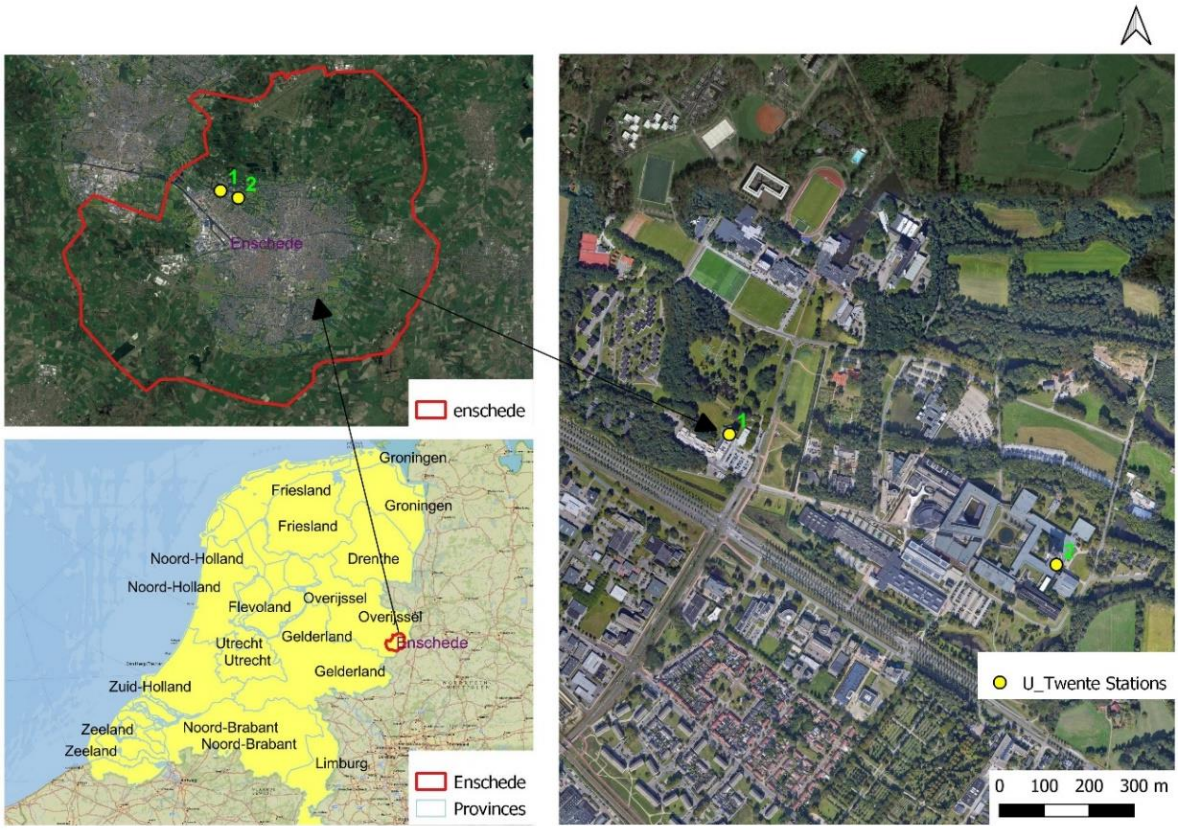


Figure 1. Location of the study area in the University of Twente. Base map: Google satellite.

2.1.1. PhenoCam stations

The two PhenoCam stations have been operational since June 25, 2019, and remain active. Table 1 presents the characteristics of the stations, and Figures 2 and 3 display views from these stations.

Table 1. PhenoCam stations. Source: PhenoCam site.

Station Characteristics	Utwente1	Utwente2
Latitude	52.240077	52.237472
Longitude	6.850103	6.86078
Elevation	25	30
Start Date	2019-06-25	2019-06-25
Location	Building de Spiegel	Building de Horst
Camera orientation	NNW	ESE
Temporal resolution	30 minutes	30 minutes

Figure 2. Station Utwente1. Source: PhenoCam site.

Figure 3. Station Utwente2. Source: PhenoCam site.

2.2. Data collection

2.2.1. PhenoCam images

PhenoCam images are three-channel (RGB) images with a temporal resolution of 30 minutes. This frequent interval allows for detailed observation of changes in vegetation over time, providing valuable insights into the dynamics of phenological processes. The images are freely available and can be downloaded through a free account on the PhenoCam website, where users can select specific timestamps and dates per station (A. Richardson, 2023). For this study, images were selected for download from March 1 to June 30, between 9:00 and 15:00 each year, resulting in 12 images per day for each station. Table 2 summarizes the images downloaded.

Table 2. Images downloaded

Year	Dates	Utwente1 images	Utwente2 images	Total images
2020	March 21 – June 30	120	121	241
2021	March 21 – June 30	117	122	239
2022	March 21 – June 30	121	121	242
2023	March 21 – June 30	116	117	233
	Total	474	481	955

2.2.2. Phenological Network Observations PEP725

The PEP725 phenological network is a comprehensive database that gathers phenological observations from across Europe. This standardized data covers various plant species and is essential for understanding the impact of climate change on plant phenology (Scheifinger, 2023). The data for this study was provided by the University obtained from the network management team.

The sources and specifications of the data used in this study are summarized in Table 3.

Table 3. Data source and specifications.

File	Source	Owner	Restrictions and license	Format	Personal data
PhenoCam images	Secondary data	PhenoCam site	No restrictions	jpg	No
Phenological networks observations	Secondary data	PEP725 site	No restrictions	csv	No

3. RESEARCH METHOD

3.1. Data preprocessing

The preprocessing of the PhenoCam images involved several steps to ensure the data were organized and of high quality. Initially, images were renamed according to the structure station_year_month_day.jpg (e.g., ut1_2020_5_20.jpg) to simplify identification and retrieval for further analysis. Subsequently, one of the twelve downloaded images per day was selected based on brightness values, prioritizing those with the highest values. Brightness is a metric utilized in various image applications (Smyth et al., 2022), enabling the identification of well-illuminated and clear images. Following this selection, a visual inspection was conducted to verify that the images were correctly selected.

Five regions of interest (patches) were identified within these images to focus on specific vegetation areas (see Figures 4 and 5). Patches 1 and 2 were selected from the images of the station UTwente1, while patches 3, 4, and 5 correspond to images from the station UTwente2. These patches were selected based on areas with deciduous vegetation, ensuring a consistent patch size of 180x180 pixels. This size was chosen to guarantee that the patches included the desired vegetation type while excluding elements such as sky, water, roads, buildings, people, and other vegetation types. The images were renamed according to the structure station_patch_year_month_day.jpg (e.g., ut2_p3_2021_4_15.jpg). Table 4 presents the number of image patches per station for the four-year period (2020-2023).

Figure 4. Selected patches from station Utwente1.

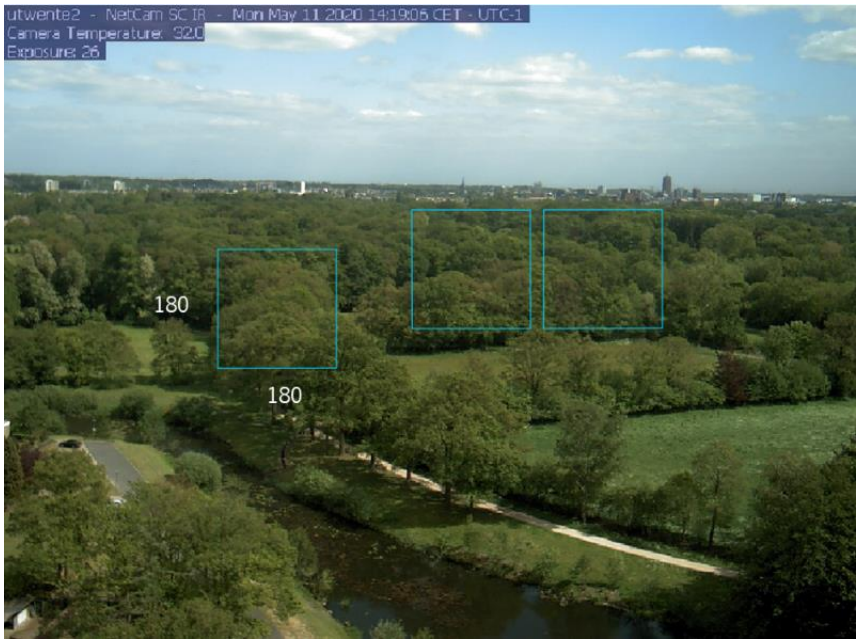


Figure 5. Selected patches from station Utwente2.

Table 4. Number of patch images per station per year.

Year	Utwente1 patches	Utwente2 patches	Total patches
2020	240	363	603
2021	234	366	600
2022	242	363	605
2023	232	351	583
Total	**948**	**1443**	**2391**

The process of determining the date to differentiate between the two phenological states involves a qualitative assessment. Visual inspection will be the primary method to identify the phenological transitions. To provide additional validation and context to this approach, a quantitative method using Green Chromatic Coordinates (GCC) will be applied. Additionally, data from PEP725's observations will be compared to support the findings.

3.1.1. Visual inspection

For this approach, I systematically reviewed each image in chronological order, meticulously identifying dates where noticeable changes in the presence of green leaves were evident. These specific dates serve as markers to distinguish between two distinct phenological stages. Table 5 shows the identified dates.

Table 5. Dates identified by visual inspection.

year	Utwente1 first leaves	Utwente2 first leaves
2020	2020-04-12	2020-04-10
2021	2021-04-25	2021-04-26
2022	2022-04-15	2022-04-18
2023	2023-04-21	2023-04-23

3.1.2. GCC time series

GCC is a colour index used to measure the relative brightness of the green channel in the given images (A. Richardson, 2023). It quantifies the amount of green colour present in a particular scene or object, making it a useful metric for analysing vegetation health, seasonal changes, and environmental conditions (A. Richardson, 2023). The GCC index is calculated using the formula:

$$GCC = \frac{G}{R + G + B}$$

The GCC approach involved generating time series data by applying this formula to the patch images. Each value within the time series represents the daily mean GCC. Subsequently, a simple moving average (SMA) filter was applied to the GCC time series to reduce noise and smooth the curve, ensuring a more accurate analysis (Bauer et al., 2019). Additionally, the GCC time series were normalized within the range of 0 to 1, promoting consistency and comparability across different datasets (Kong et al., 2022). To determine the date to differentiate between the two phenological states and facilitate dataset classification, two thresholding methods were utilized: one set at 0.3 and another at 0.5. These threshold values are commonly used in studies related to vegetation phenology, such as the study by Kong et al. (2022). The identified date corresponds to the point where the normalized time series intersects the threshold value (Kong et al., 2022). The dates identified by the two threshold methods are displayed in Table 6 and figures 6 and 7.

Table 6. Dates identified by the two threshold methods.

trs_0.3		trs_0.5	
Utwente1	Utwente2	Utwente1	Utwente2
2020-04-16	2020-04-13	2020-04-22	2020-04-22
2021-05-06	2021-05-01	2021-05-10	2021-05-07
2022-04-20	2022-04-17	2022-04-23	2022-04-21
2023-04-27	2023-04-18	2023-05-02	2023-04-24

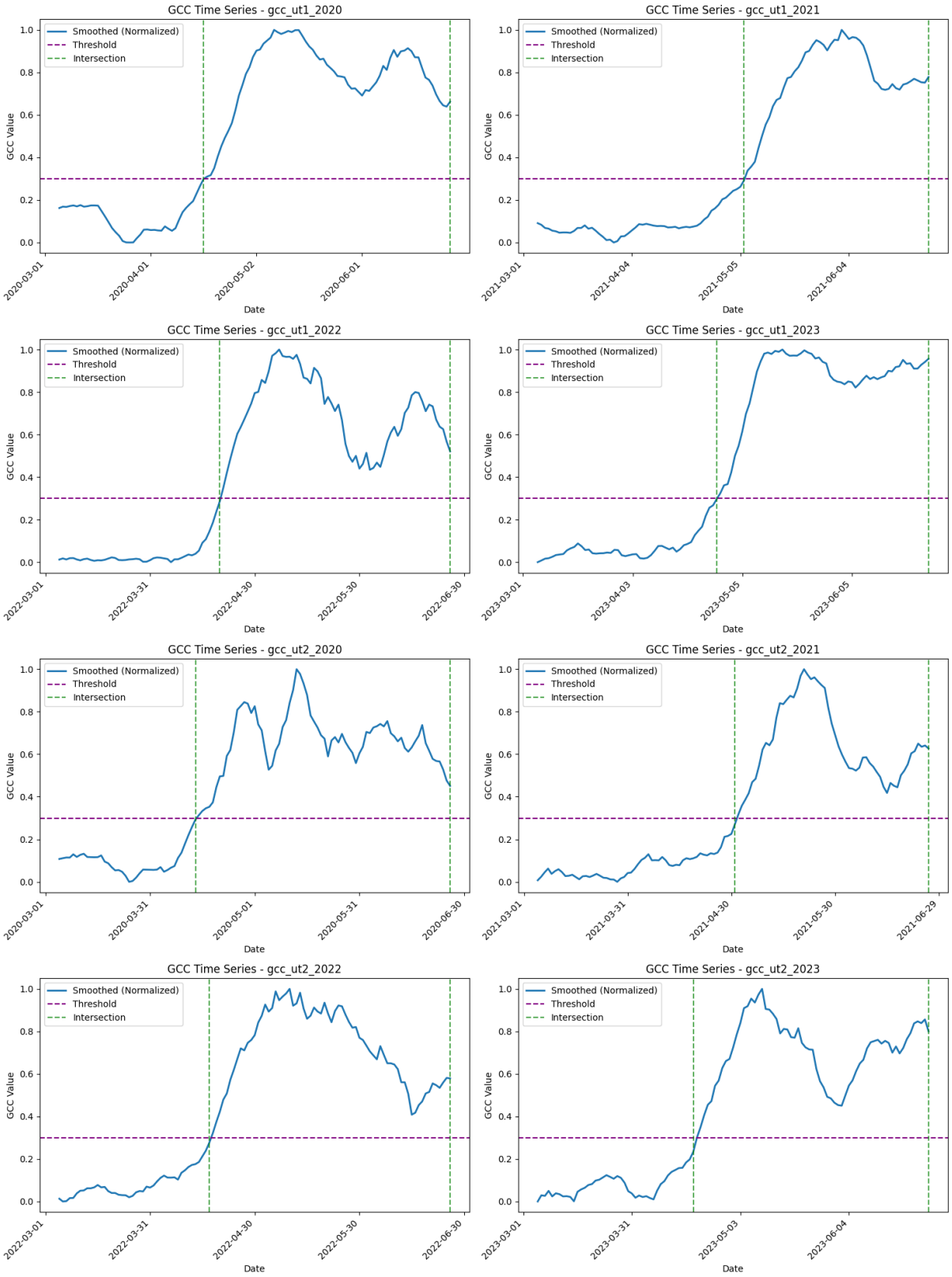


Figure 6. GCC time series for threshold 0.3

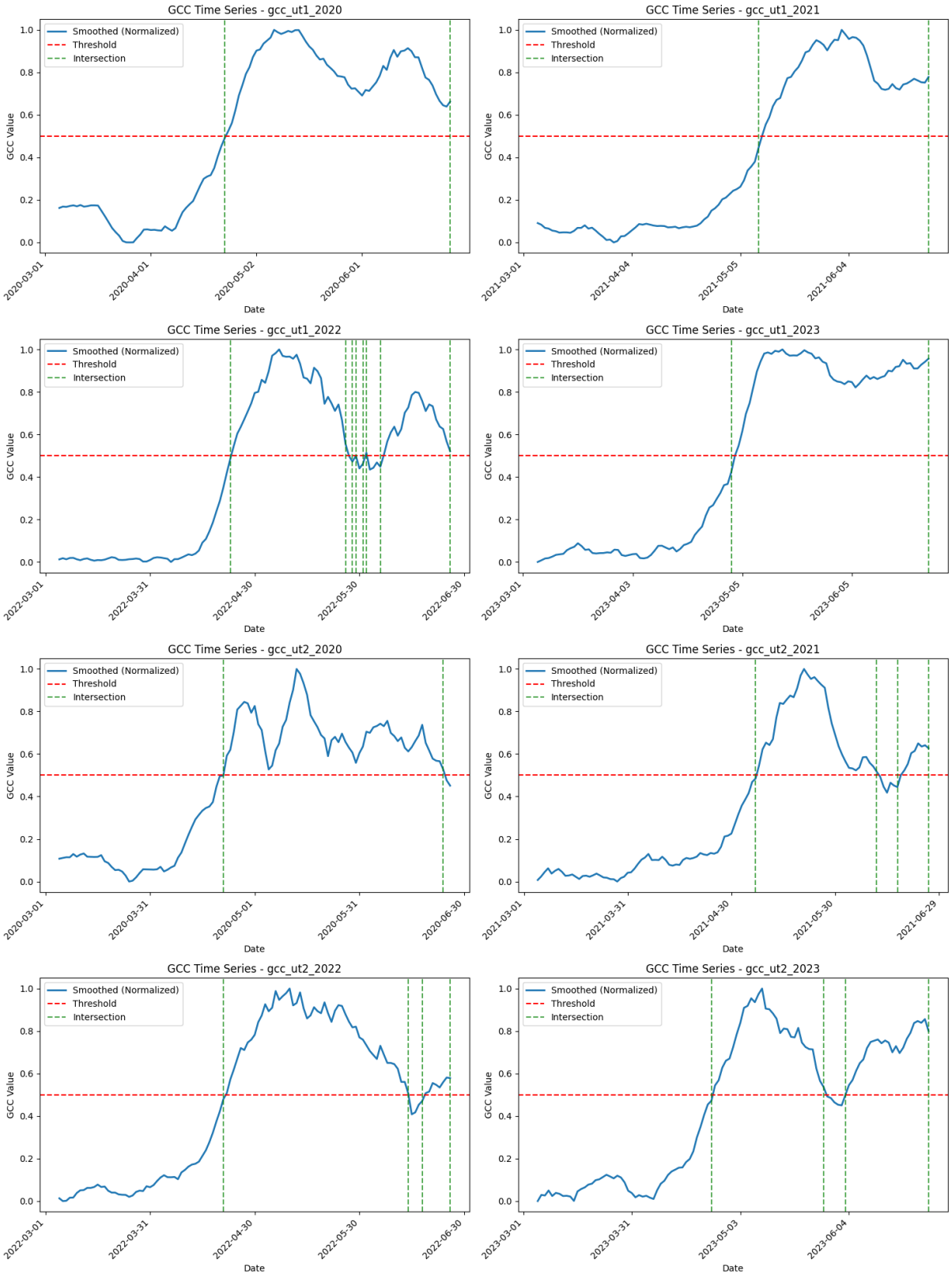


Figure 7. GCC time series for threshold 0.5

3.1.3. PEP725

The PEP725 approach focused on selecting the closest observation points available in Germany, given the absence of observation points in the Netherlands. These selected points were approximately 19 kilometers away from the study area (see Figure 8). To ensure relevance to the study, a filter was applied to specifically target the genus Quercus, which is the same genus present in the study area. Additionally, a second filter was implemented to select phase ID 11. This phenological phase corresponds to 'first true leaf, leaf pair or whorl unfolded, P: First leaves unfolded' in the BBCH-scale, a standardized system for identifying plant developmental stages. The date associated with this phenological phase was selected as the reference to differentiate between the two phenological states. It is important to note that information for this phase and genus was only available for the years 2020 and 2021 (see Table 7).

Table 7. Dates identified using PE725.

s_id	lon	lat	alt	genus	species	phase_id	year	date
934	7.11667	52.3	50	Quercus	Quercus robur	11	2020	2020-04-18
934	7.11667	52.3	50	Quercus	Quercus robur	11	2021	2021-05-10

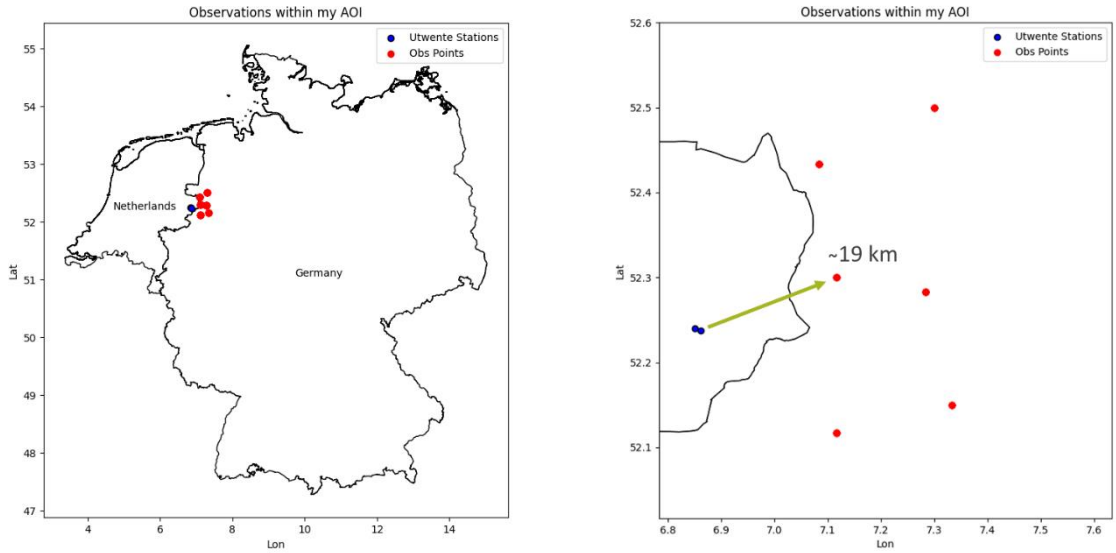


Figure 8. Closest observations point to the study area stations. Source: PEP725.

3.2. Labelling

For the labelling process, the dates determined by the four different approaches were compared (see Table 8). It is important to note that these dates showed close agreement, indicating coherence across methods, and reinforcing the reliability of visual inspection in determining the phenological stages.

Table 8. Comparison of approaches for phenological state differentiation.

year	Pep725	Visual inspection		trs30		trs50	
		utwente1	utwente2	Utwente1	Utwente2	Utwente1	Utwente2
2020	2020-04-18	2020-04-12	2020-04-10	2020-04-16	2020-04-13	2020-04-22	2020-04-22
2021	2021-05-10	2021-04-25	2021-04-26	2021-05-06	2021-05-01	2021-05-10	2021-05-07
2022		2022-04-15	2022-04-18	2022-04-20	2022-04-17	2022-04-23	2022-04-21
2023		2023-04-21	2023-04-23	2023-04-27	2023-04-18	2023-05-02	2023-04-24

Table 9 provides the start and end dates for each phenological class, along with its corresponding number of days.

Table 9. Start and end dates of phenological classes.

Year	class_1			class_2		
	Beginning	End	Days	Beginning	End	Days
2020	2020-03-01	2020-04-12	43	2020-04-13	2020-05-25	42
2021	2021-03-01	2021-04-25	56	2021-04-26	2021-06-15	51
2022	2022-03-01	2022-04-15	46	2022-04-16	2022-05-31	46
2023	2023-03-01	2023-04-21	51	2023-04-22	2023-06-12	49

3.3. Convolutional Neural Network

Dataset splitting ensures that the CNN model is trained on a representative subset of data, validated for parameter optimization, and tested on an independent set to evaluate generalization (Kattenborn et al., 2021). Following this approach, the dataset is divided as follow (see table 10):

Table 10. Dataset distribution.

Dataset	Year	Class 1 patches	Class 2 patches	Total patches
Training	2020	213	210	423
	2021	270	255	525
Validation	2022	230	230	460
Testing	2023	251	242	493

For this setup, the model is trained to learn from historical patterns and the validation set helps in fine-tuning the model's parameters for optimal performance on unseen data. Finally, the testing dataset evaluates the model's ability to generalize to future scenarios(Ferchichi et al., 2024).

A CNN model is composed of several types of layers, each serving a specific function. To build a CNN model involves choosing appropriate layers, starting with the convolutional layers, which are essential for learning spatial hierarchies, and followed by pooling layers that help spatial dimensions. Fully connected layers at the end process the learned features for classification (Kattenborn et al., 2021) (see Figure 9).

To effectively implement these layers, it is important to understand their roles and the activation functions that help the model learn from data. A suitable activation function such as Rectified Linear Unit (ReLU) introduces non-linearity, allowing the model to learn the complexity of the data during convolutional and initial dense layers (Kattenborn et al., 2021). The last dense layer will use a softmax activation for classification, representing the likelihood or probability of each class. For pooling layer, the network focuses on reducing the spatial dimension of the input by downsampling the information, characterized by the pooling size (pooling window) and the step size by which the window moves (Kattenborn et al., 2021). Fully connected layers transform the features learned by convolutional and pooling layer into a suitable format for making class predictions (Kattenborn et al., 2021). The last dense layer will have two neurons representing the two phenological states.

The network's output will be a 1D vector with class probabilities for each image, where the model classifies based on the highest probability among classes. Key considerations include using dropout rates to prevent

overfitting and fine-tuning model parameters to ensure accurate classification of phenological stages (Kattenborn et al., 2021).

Designing the CNN model is an iterative process and involves the refinement of hyperparameters (Kattenborn et al., 2021) such as:

- Type of model: sequential, layers are added sequentially.
- Number of layers: Task-dependent, number of layers can vary.
- Number of filters: Determines how many features the convolutional layer can recognize from the input data.
- Kernel size: Specifies the size of the filter used in convolutional layers.
- Number of units or neurons per layer: Determines the dimensionality of the fully connected layer.
- Dropout rate: Randomly ignores a fraction of units during training to prevent overfitting.
- Epochs: number of times the entire dataset is passed through the neural network during training.
- Optimizer: A specific algorithm used to update the weights of the neural network during training to minimize the loss function (Pannu et al., 2020).
- Batch size: number of training samples utilized in one iteration.
- Learning rate: It controls the step size or rate at which the model learns.

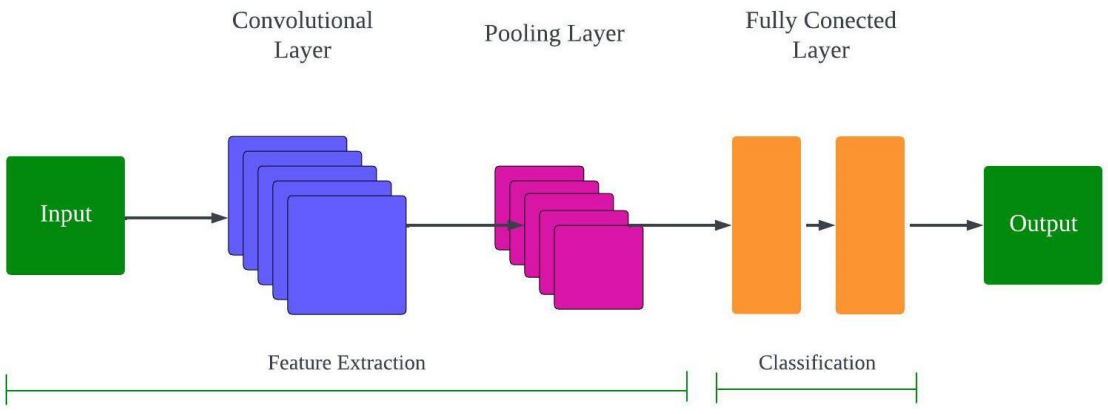


Figure 9. CNN architecture.

3.4. Random Forest

To provide a comparative analysis, a Random Forest (RF) classifier is implemented alongside the CNN model. The Random Forest algorithm, known for its robustness and ability to handle high-dimensional data, is an ensemble learning method that constructs multiple decision trees during training for classification tasks (Li et al., 2021). The number of trees and the maximum number of nodes in each tree are critical hyperparameters of RF models. (Li et al., 2021)

For the RF model, the dataset was structured differently compared to the CNN input. Instead of using the 180x180x3 RGB images, a set of engineered features was implemented, consisting of 10 features designed specifically for capturing various aspects of vegetation characteristics:

- Mean RGB values (3): GCLM-based texture features derived from the average intensity across the red, green, and blue bands (Zhang et al., 2021).
- Variance per band (3): GCLM-based texture features that quantify the variability of pixel intensities across the red, green, and blue bands (Zhang et al., 2021).
- Standard deviation per band (3): Texture features that complement variance by indicating the spread of pixel intensities within each band.
- GCC (1): Green Chromatic Coordinate, a vegetation index that measures the relative greenness of vegetation (A. Richardson, 2023).

This transformation into a 1x1x10 format aimed to capture essential statistical information about the images while reducing the dimensionality, making it suitable for the RF model.

3.5. Performance assessment

The performance of the CNN model is evaluated during both the validation and testing phases using accuracy as a fundamental metric, which represents the proportion of correctly classified instances out of the total instances, for measuring the model's effectiveness in correctly classifying phenological states (Kattenborn et al., 2021). During the validation phase, the model's accuracy is calculated to ensure proper learning from the training data, crucial for fine-tuning model parameters and identifying potential overfitting or underfitting issues (Kattenborn et al., 2021). After model fine-tuning and validation, the model's performance is assessed using the testing dataset provide an unbiased estimate of its ability to generalize to unseen data (Kattenborn et al., 2021).

3.5.1. Other metrics

Other common metrics are applied to evaluate the models (Kattenborn et al., 2021):

- F1 Score: Mean of precision and recall.
- Precision: Measures the accuracy of positive predictions.
- Recall: Ratio of true positive predictions to the total actual positive instances.

3.5.2. Learning curve

The learning curve is a crucial tool for understanding the performance of the CNN model over the training and validation phases. During the training phase, it tracks changes in the model's accuracy and loss metrics across epochs. In the validation phase, the learning curve assesses the model's performance on unseen data at regular intervals. Analysing the validation performance enables adjustments to the model's architecture and hyperparameters to enhance overall model performance and prevent overfitting (Kattenborn et al., 2021).

The Random Forest model's performance is evaluated through accuracy metrics and learning curves (Li et al., 2021). Accuracy is calculated during both validation and testing phases to measure the model's effectiveness in correctly classifying phenological states. The learning curve, showing accuracy over the number of trees is analysed to understand the model's learning dynamics and its ability to generalize. This analysis helps in fine-tuning of hyperparameters to enhance overall performance and prevent overfitting.

4. RESULTS AND DISCUSSION

4.1. CNN

4.1.1. Model architecture

To define an optimal architecture for the CNN model, I carefully reviewed and based the design on several studies that implemented similar architectures for image classification tasks (Kattenborn et al., 2021; Pannu et al., 2020). Using TensorFlow and Keras as powerful deep learning frameworks, I adopted a foundational design of two convolutional layers with ReLU activation functions. These were followed by max pooling layers to progressively reduce spatial dimensions and extract relevant features. Dropout layers were added to prevent overfitting, and dense layers for the binary classification purpose.

This architecture can be described as follows:

1. First convolutional layer: Includes 32 filters of size 3x3, applying ReLU activation, and accepts an input shape of 180x180x3.
2. First max pooling layer: Performs max pooling with a pool size of 2x2.
3. Second convolutional layer: Includes 64 filters of size 3x3 with ReLU activation.
4. Second max pooling layer: Performs max pooling with a pool size of 2x2.
5. Dropout layer (25%): Introduces dropout regularization to reduce overfitting.
6. Flatten layer: Transforms the output from the convolutional layers into a vector for the fully connected layers.
7. First dense layer (128 neurons): Fully connected layer with ReLU activation.
8. Dropout layer (50%): Further applies dropout regularization.

9. Output layer: Final layer with 2 neurons and softmax activation for classification.

4.1.2. Model summary

The model is defined in a sequential mode, meaning that layers are added in sequence. Based on the model´s architecture, Table 11 shows the output shape of the tensor after passing through each layer and the number of parameters (weights) associated with each layer, these parameters are learned during the training process. The model is illustrated in Figure 10.

Table 11. Model summary.

Model: "sequential"		
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 178, 178, 32)	896
max_pooling2d (MaxPooling2D)	(None, 89, 89, 32)	0
conv2d_1 (Conv2D)	(None, 87, 87, 32)	18496
max_pooling2d_1 (MaxPoolingd2D)	(None, 43, 43, 32)	0
dropout (Dropout)	(None, 43, 43, 32)	0
flatten (Flatten)	(None, 59168)	0
dense (Dense)	(None, 128)	15147136
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 2)	258
Total params: 15166786 (57.86 MB)		
Trainable params: 15166786 (55.86 MB)		
Non-trainable params: 0 (0.00 Byte)		

The detailed summary offers insights into how the model processes data. This structural overview is fundamental for understanding the model's architecture and its computational requirements during training and predictions on new data.

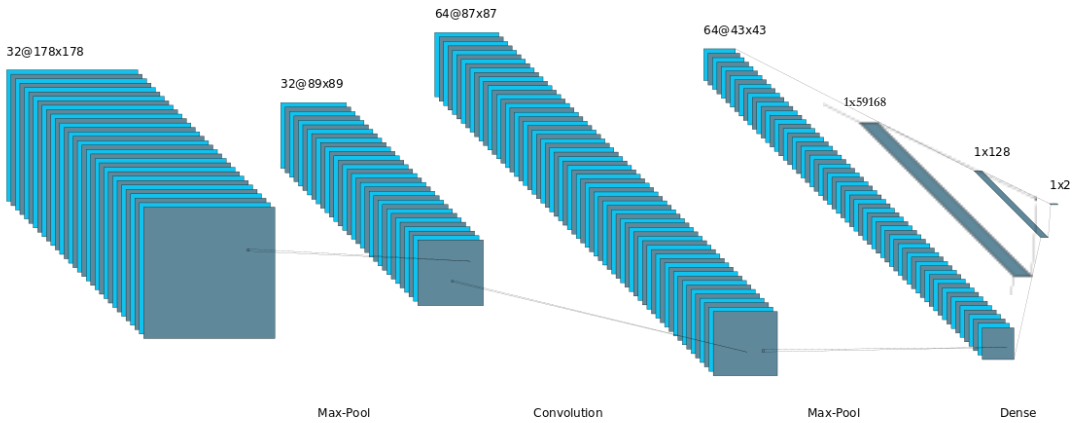


Figure 10. CNN model architecture

4.1.3. Hyperparameters

The hyperparameter values play a crucial role in determining the model's performance during training, as previously detailed. To find the optimal model's hyperparameters, I considered several factors.

Firstly, early stopping was implemented with a patience of 10 epochs, this is a functionality provided by Keras. This mechanism helps in preventing overfitting by terminating the training when the validation loss stabilizes. A patience of 10 epochs means that the training process will end if there is no improvement in the validation loss for 10 consecutive epochs, ensuring the model does not continue to train on fluctuations, thereby avoiding premature stopping before reaching optimal performance.

The optimizer is critical in training neural networks as it adjusts the learning rate and updates the model weights to minimize the loss function (Kattenborn et al., 2021). Initially, I used the Stochastic Gradient Descent (SGD) optimizer, incorporating a momentum of 0.9 as implemented by Pannu et al. (2020), and set the learning rate to 0.01 over 50 epochs. The purpose of using momentum is to accelerate SGD for faster convergence. However, this configuration produced unsatisfactory results, with an accuracy of 0.5 and a loss of 0.693. This outcome indicated that the model was not learning effectively.

After analysing these initial results, I adjusted the optimizer to plain SGD while keeping the same learning rate and epoch settings. This change led to a significant improvement in the model's performance, achieving an accuracy of 0.945 and a loss of 0.122 (see Figure 11). This substantial improvement suggests that the simpler optimization strategy allowed for more stable and effective learning in this context.

Table 12 shows the hyperparameter values during fine-tuning.

Table 12. Hyperparameter values

Hyperparameter		
Optimizer	SDG	SGD
Learning rate	0.01	0.01
Batch size	32	32
Epochs	50	50
Early stopping patience	10	10
Momentum	0.9	-

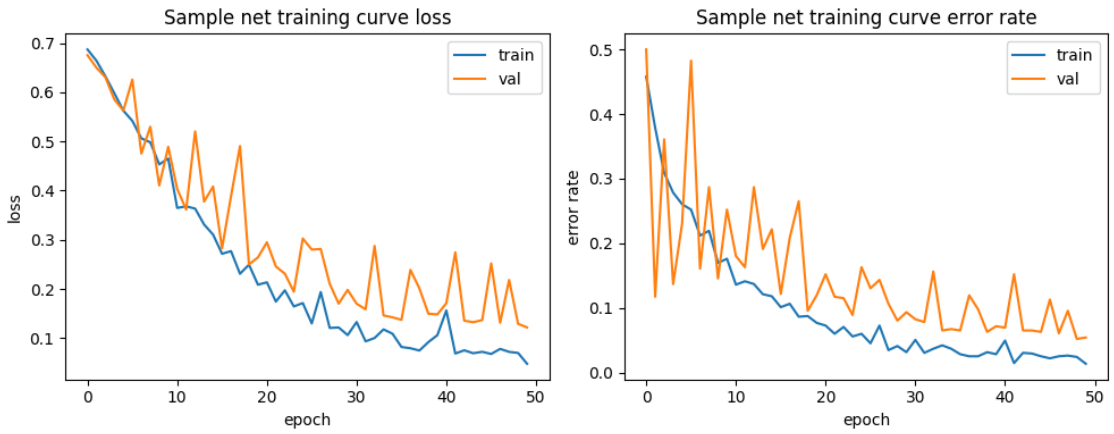


Figure 11. Loss and error curve of the CNN model during the training and validation phase.

4.1.4. Intermediate results

Intermediate results are visualized to gain deeper insight into the CNN model (see Figure 12). These visualizations include the input images with their corresponding class labels, the first convolutional layer where low-level features like texture are captured, followed by the max-pooling layer that reduce spatial dimensions. Subsequently, the visualization progress through the second convolutional layer, the second max-pooling layer and finally the first dense layer.

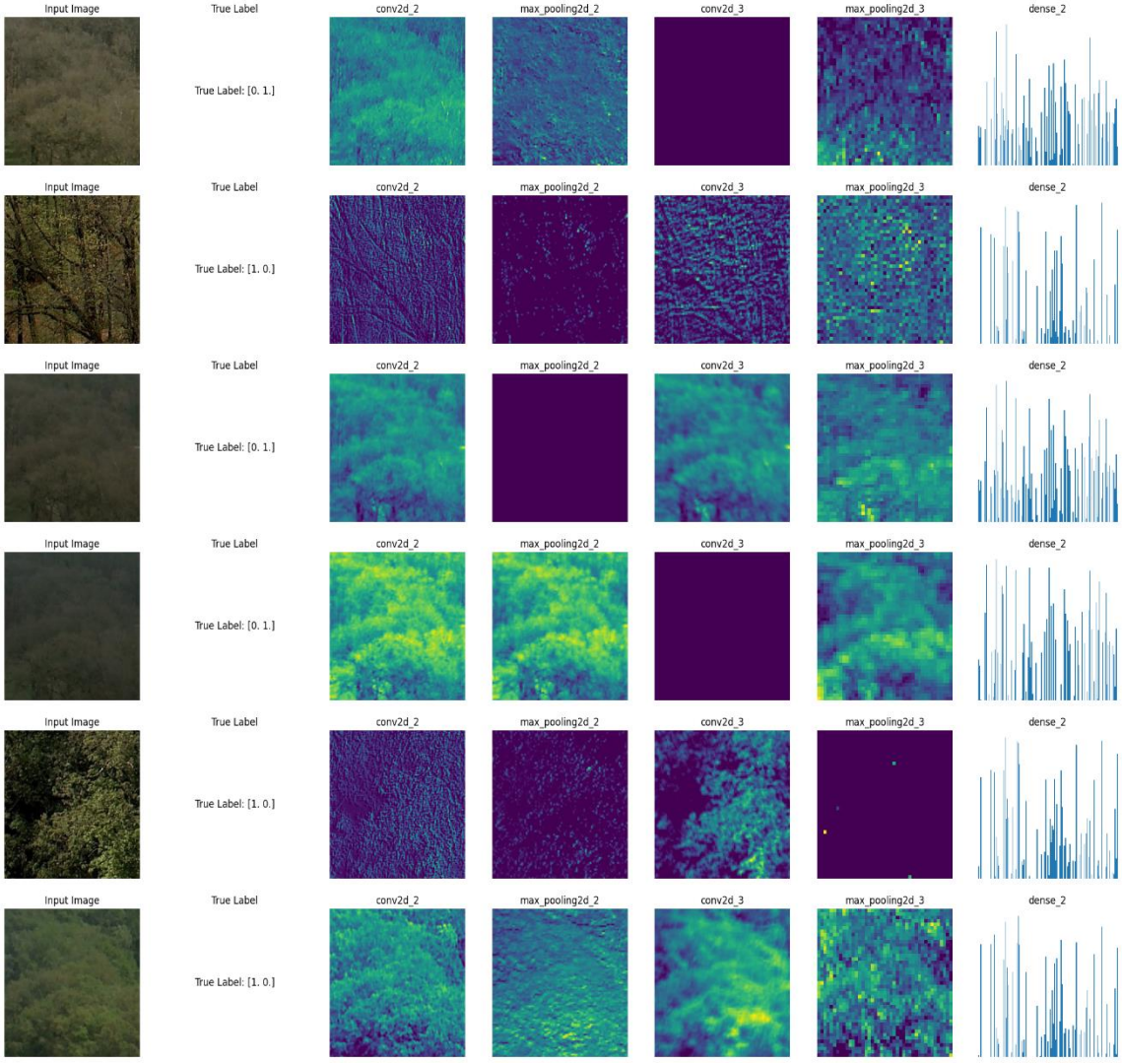


Figure 12. Intermediate results for the CNN model during the testing phase.

After going through these intermediate outputs, it is clearer how the model processes and interprets the extracted features before making final predictions. The first convolutional layer identified fundamental elements for distinguishing between spring green up and summer green peak. In the case of spring green up, the layer highlighted details like branches, typical characteristic of early growth in vegetation before the emergence of leaves. Conversely, during summer green peak, the layer emphasized the textured patterns of rich foliage, indicative of dense vegetation. This feature extraction process effectively captured the distinctive characteristics associated with the seasonal transition.

Moreover, maxpooling layers play a critical role, by downsampling feature maps generated by convolutional layers. Comparing a maxpooling layer to its preceding convolutional layer shows a noticeable the reduction in details; for some samples this layer presents almost no discernible features. maxpooling achieves this by smoothing out finer textures and patterns present in the convolutional layer's output, promoting spatial invariance in subsequent layers for higher-level processing within the network. In contrast, the output images from the second convolutional layer reveal more details compared to the preceding maxpooling layer. Notably, the use of a colormap exposes a variety of colours, indicating that the second convolutional layer captures a diverse range of features. This diversity highlights the layer's ability to recognize complex patterns and textures in the input data.

Furthermore, the bar plot visualization of the dense layer reveals distinct patterns between spring green up and summer green peak classes. For summer green peak, the bars exhibit a consistent density across the entire series, suggesting a uniform distribution of neuron activations or weights. This consistency indicates that the model's decision-making process for identifying summer green peak relies on a balanced contribution from various neurons throughout the dense layer. In contrast, the bars for spring green up appear more dispersed, with higher values observed in the middle sections of the plot. This dispersion suggests that certain neurons in the dense layer are more active or weighted more heavily at specific points, reflecting the features and patterns characteristic of the vegetation during the spring green up.

4.1.5. Performance assessment

4.1.5.1. Testing

The accuracy on the testing set serves as a crucial metric for evaluating the overall performance of the trained model, as shown in Table 13. It measures the proportion of correctly classified instances among the total instances in the testing set, providing an indication of how well the model generalizes to new, unseen data.

To gain a deeper understanding of the CNN model's performance, a confusion matrix was generated (see Figure 13). This matrix offers a detailed overview of how test instances were classified, indicating the classification accuracy and misclassification instances. In this context, class 0 corresponds to spring green up, and class 1 to summer green peak. The confusion matrix allows for the examination of true positives, true negatives, false positives, and false negatives, providing insights into specific areas where the model may struggle.

Table 13. Loss and accuracy of the CNN model during the testing phase

Test Loss	Test Accuracy
0.085	0.974

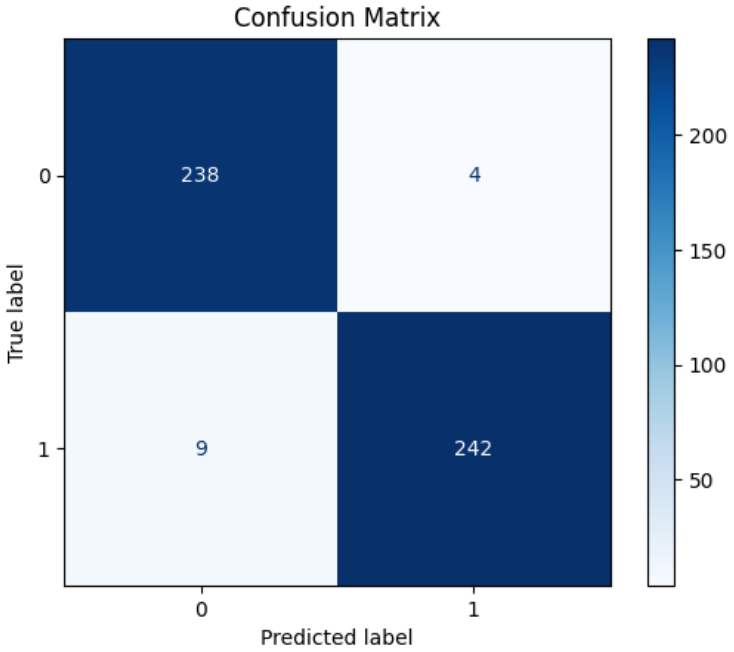


Figure 13. CNN model confusion matrix

4.2. Random Forest

The random forest model has been implemented for comparison against the CNN model. The validation accuracy provides insight into how well the model generalizes during the training phase, while the test accuracy indicates its performance on unseen data.

The parameter values for the Random Forest model were selected through an iterative process aimed at optimizing the classifier's performance. This involved adjusting parameters such as the number of trees and tree depths, and evaluating performance metrics, including accuracy, across both the training and validation phases (see Figure 14). Initially, varying the number of trees and depths showed limited impact on the model's accuracy, which remained relatively stable across different configurations (see Table 14).

Table 14. Validation accuracy with different parameters values

n_trees	Max_depth	Val_accuracy
50	20	0.930
25	10	0.939
20	10	0.935
20	7	0.935

Figure 14. Training and validation accuracy curve for RF classifier.

After fine-tuning the parameters, the Random Forest model was defined as follows:

1. n_estimators: twenty-five decision trees.

2. max_depth: the maximum depth of each decision tree is ten.

The model demonstrated robust performance on the test set, achieving an accuracy of 0.965, which suggests strong generalization to unseen data. To gain a better understanding of its performance, a confusion matrix was generated (see Figure 15). This matrix provides a detailed overview of how test instances were classified, revealing only a few misclassifications. In this context, class 0 corresponds to spring green up, and class 1 corresponds to summer green peak.

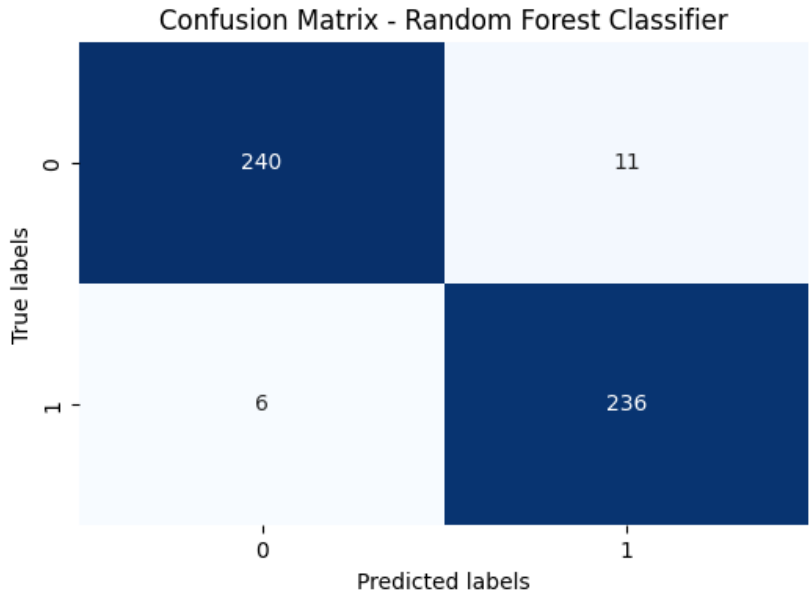


Figure 15. Random Forest confusion matrix.

Feature importance analysis was also applied to the Random Forest model (see Figure 16). It was observed that the GCC feature holds the highest importance, with a value of approximately 0.7, indicating its significant role in distinguishing between the phenological stages. Additionally, the green-related features, variance, and standard deviation of the green band, ranked as the second and third most important features, respectively. Despite their importance, their values were still below 0.1, indicating that their individual contributions are relatively minor compared to GCC. The remaining features displayed even lower importance values.

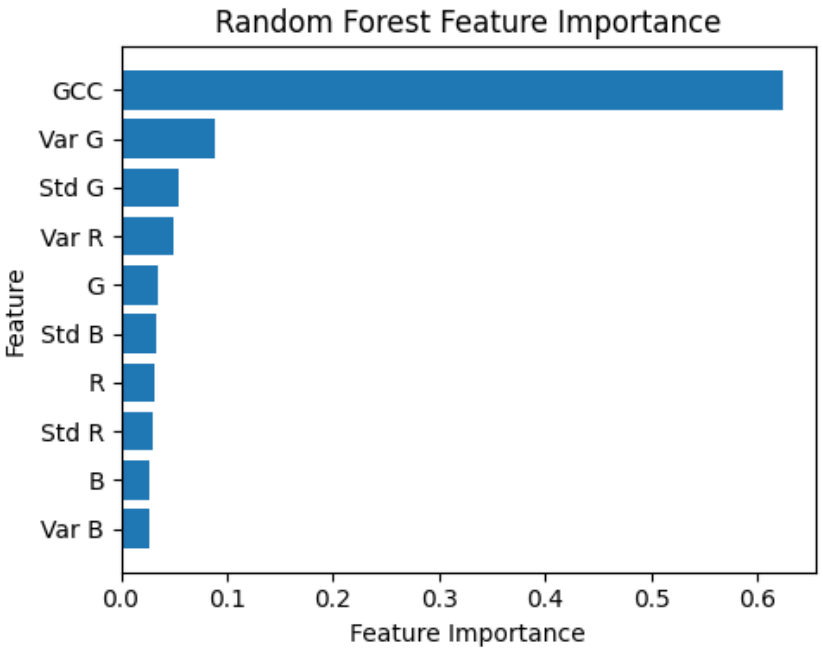


Figure 16. Random forest feature importance.

4.3. Comparison of models

The learning curves, particularly in terms of accuracy, provide valuable insights into the training dynamics and generalization capabilities of both models. The CNN's accuracy improves gradually at a slower rate, indicating ongoing refinement in capturing important patterns and features for distinguishing phenological stages. In contrast, the RF model's accuracy curve stabilizes early in training, demonstrating its robustness and rapid convergence to high accuracy. The CNN's learning curve suggests deeper feature extraction and learning capacity.

Overall, both models exhibited high performance across key evaluation metrics, highlighting their efficacy in classifying phenological stages (see Table 15). Despite their distinct approaches, with RF relying on feature engineering and CNN learning complex patterns, the CNN model demonstrated slightly superior performance. It achieved marginally higher precision, indicating fewer incorrect positive predictions, and a slightly higher F1 score, reflecting a better balance between precision and recall. Both models showed strong recall rates and high accuracy.

Table 15. Evaluation metrics for both models.

Evaluation metric	RF	CNN
Precision	0.95	0.97
Recall	0.97	0.97
RF1	0.96	0.97
Accuracy	0.96	0.97

When comparing the confusion matrices of the CNN and RF models, both demonstrate strong performance in classifying spring green-up (class 0) and summer green peak (class 1). The CNN exhibits superior accuracy in classifying class 1, achieving higher true positives and fewer misclassifications compared to the RF model. In contrast, the RF model shows stronger results in classifying class 0, with fewer instances misclassified as class 1 compared to the CNN.

A critical distinction between the models lies in their dataset structures and their implications for robustness and generalization. The RF model operates on a simpler dataset structure of 1x1x10, where each instance is represented by a flat vector of ten features. In contrast, the CNN model processes data in a more complex structure of 180x180x3, where each instance is represented by a 180x180 pixel image with three color channels. This higher-dimensional input allows the CNN to perform feature extraction, learning spatial and spectral patterns inherent in the data. Consequently, while the RF model exhibited stability and strong performance with predefined features, the CNN's capacity to learn patterns and features from data suggests it may generalize more effectively to new, unseen data, making it adaptable to different environmental conditions and datasets.

5. CONCLUSION

This study developed a CNN model for binary classification using PhenoCam images to accurately classify key phenological stages, providing valuable insights into the seasonal dynamics of deciduous trees. The study contributes to phenological research by offering a robust framework for predicting vegetation phenological stages. The answers to the research questions are addressed as follows:

Subobjective 1

To gather and ensure the quality of PhenoCam images spanning the four-year period (2020-2023) for the study area, along with phenological data from additional resources.

Q1. How can PhenoCam images spanning the four-year period (2020-2023) be effectively gathered and organized alongside supplementary phenological data?

PhenoCam images spanning the four-year period from 2020 to 2023 were effectively gathered and organized alongside supplementary phenological data using a systematic approach. Initially, the data was freely downloaded from the official PhenoCam website, where images were conveniently named based on station identifiers and corresponding dates. Careful selection criteria were applied to ensure the dataset's quality for subsequent processing steps. Additionally, supplementary data from PEP725, filtered based on vegetation type, phenological phase, and proximity to the study area, was integrated to enhance the phenological context and accuracy of the dataset.

To assign labels to the images based on their respective phenological states.

Q2. Which approaches and additional datasets could be utilized to label PhenoCam images based on their phenological states?

Visual inspection served as the primary method to differentiate between the two phenological states due to its traditional and practical approach, while the threshold methods and data from PEP725 provided additional validation and support.

The dataset was structured into training, validation, and test sets based on the years: data from 2020 and 2021 were used for training to capture historical trends in phenological patterns. Data from 2022 were set for validation to fine-tune model parameters and optimize performance. Finally, data from 2023 were reserved for testing the model's predictive capabilities on unseen data from the latest year.

Subobjective 2

To develop a CNN model that identifies phenological states for deciduous trees, specifically spring green-up and summer green peak, for a four-year period (2020-2023) by classifying images into these two key phenological states.

Q3. To what extent does the choice of architecture and parameters for the CNN model influence its ability to identify the phenological states?

The choice of CNN architecture and parameters significantly influences the model's ability to accurately identify phenological states. Based on a thorough literature review focusing on CNN applications in classification tasks, the model's architecture was selected to emphasize feature extraction and hierarchical learning. Fine-tuning of hyperparameters played a crucial role in developing a robust model capable of distinguishing between spring green-up and summer green peak effectively.

Furthermore, key considerations included monitoring the learning curve to ensure optimal performance during the training phase. Techniques such as early stopping were implemented to prevent overfitting, while adjustments to the learning rate and filter configurations were made to enhance feature learning and model convergence. These strategies contributed to the development of a CNN model with high accuracy and generalization capability in classifying phenological states from PhenoCam images.

Subobjective 3

To assess the performance of the developed CNN model in identifying phenological states for deciduous trees and compare it with traditional machine learning models.

Q4. How does the developed CNN model identify the phenological states for deciduous trees?

The developed CNN model employs a deep learning architecture designed to extract and learn hierarchical features from PhenoCam images. It identified phenological states by analysing patterns and textures captured in the images.

Q5. How does the CNN model’s performance in identifying the phenological states compare with traditional machine learning models?

The CNN model demonstrates high performance in identifying phenological states compared to a traditional machine learning model (RF). The CNN model showed slightly higher performance across all evaluated metrics. This superiority can be attributed to the CNN's capacity for automatic feature extraction from image data, which allows it to learn patterns and features that traditional models may struggle to capture with engineered features alone.

5.1. Limitations

Firstly, the dataset spans only four years, limiting the understanding on how phenological stages of deciduous trees may have evolved over a longer timeframe. Additionally, the study focused on a limited number of PhenoCam; including more stations across a broader area with similar vegetation types could enhance the robustness of the analysis. While PEP725 provided valuable data, its lack of availability for the Netherlands restricted the scope of supplementary information. Furthermore, due to limited expertise in feature engineering, the study may not have fully leveraged all potential features for the RF model. Finally, there is a lack of studies utilizing PhenoCam images for classification tasks to use as reference.

5.2. Future work

Addressing the limitations mentioned before could open new paths for future research. Extending the timeframe of the dataset and expanding the number of PhenoCam stations across diverse geographic locations with similar vegetation types would enhance phenological research. Additionally, incorporating a temporal dimension to benefit from the high-frequency images captured by PhenoCams would provide deeper insights into the variability of phenological dynamics. Future research could also focus on exploring ensemble methods or integrating additional environmental variables to further enhance robustness in phenological studies.

LIST OF REFERENCES

Bauer, J., Jarmer, T., Schittenhelm, S., Siegmann, B., & Aschenbruck, N. (2019). Processing and filtering of leaf area index time series assessed by in-situ wireless sensor networks. *Computers and Electronics in Agriculture*, *165*. https://doi.org/10.1016/j.compag.2019.104867

Capinha, C., Ceia-Hasse, A., Kramer, A. M., & Meijer, C. (2021). Deep learning for supervised classification of temporal data in ecology. *Ecological Informatics*, *61*. https://doi.org/10.1016/j.ecoinf.2021.101252

Ferchichi, A., Chihaoui, M., & Ferchichi, A. (2024). Spatio-temporal modeling of climate change impacts on drought forecast using Generative Adversarial Network: A case study in Africa. *Expert Systems with Applications*, *238*. https://doi.org/10.1016/j.eswa.2023.122211

Ge, W., Li, X., Jing, L., Han, J., & Wang, F. (2023). Monitoring canopy-scale autumn leaf phenology at fine-scale using unmanned aerial vehicle (UAV) photography. *Agricultural and Forest Meteorology*, *332*. https://doi.org/10.1016/j.agrformet.2023.109372

Gričar, J., Jevšenak, J., Hafner, P., Prislan, P., Ferlan, M., Lavrič, M., Vodnik, D., & Eler, K. (2022). Climatic regulation of leaf and cambial phenology in Quercus pubescens: Their interlinkage and impact on xylem and phloem conduits. *Science of the Total Environment*, *802*. https://doi.org/10.1016/j.scitotenv.2021.149968

Guo, Y., Fu, Y., Hao, F., Zhang, X., Wu, W., Jin, X., Robin Bryant, C., & Senthilnath, J. (2021). Integrated phenology and climate in rice yields prediction using machine learning methods. *Ecological Indicators*, *120*. https://doi.org/10.1016/j.ecolind.2020.106935

Katal, N., Rzanny, M., Mäder, P., & Wäldchen, J. (2022). Deep Learning in Plant Phenological Research: A Systematic Literature Review. In *Frontiers in Plant Science* (Vol. 13). Frontiers Media S.A. https://doi.org/10.3389/fpls.2022.805738

Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 173, pp. 24–49). Elsevier B.V. https://doi.org/10.1016/j.isprsjprs.2020.12.010

Kong, D., McVicar, T. R., Xiao, M., Zhang, Y., Peña-Arancibia, J. L., Filippa, G., Xie, Y., & Gu, X. (2022). phenofit: An R package for extracting vegetation phenology from time series remote sensing. *Methods in Ecology and Evolution*, *13*(7), 1508–1527. https://doi.org/10.1111/2041-210X.13870

Kun, X., Wei, W., Sun, Y., Wang, Y., & Xin, Q. (2023). Mapping fine-spatial-resolution vegetation spring phenology from individual Landsat images using a convolutional neural network. *International Journal of Remote Sensing*, *44*(9), 3059–3081. https://doi.org/10.1080/01431161.2023.2216846

Lee, C. K. F., Song, G., Muller-Landau, H. C., Wu, S., Wright, S. J., Cushman, K. C., Araujo, R. F., Bohlman, S., Zhao, Y., Lin, Z., Sun, Z., Cheng, P. C. Y., Ng, M. K. P., & Wu, J. (2023). Cost-effective and accurate monitoring of flowering across multiple tropical tree species over two years with a time series of high-resolution drone imagery and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *201*, 92–103. https://doi.org/10.1016/j.isprsjprs.2023.05.022

Li, R., Xu, M., Chen, Z., Gao, B., Cai, J., Shen, F., He, X., Zhuang, Y., & Chen, D. (2021). Phenology-based classification of crop species and rotation types using fused MODIS and Landsat data: The comparison of a random-forest-based model and a decision-rule-based model. *Soil and Tillage Research*, *206*. https://doi.org/10.1016/j.still.2020.104838

Lou, Z., Wang, F., Peng, D., Zhang, X., Xu, J., Zhu, X., Wang, Y., Shi, Z., Yu, L., Liu, G., Xie, Q., & Dou, C. (2023). Combining shape and crop models to detect soybean growth stages. *Remote Sensing of Environment*, *298*. https://doi.org/10.1016/j.rse.2023.113827

Pannu, H. S., Ahuja, S., Dang, N., Soni, S., & Malhi, A. K. (2020). Deep learning based image classification for intestinal hemorrhage. *Multimedia Tools and Applications*, *79*(29–30), 21941–21966. https://doi.org/10.1007/s11042-020-08905-7

Richardson, A. (2023). *How to access PhenoCam data and imagery: A tutorial*. https://phenocam.nau.edu/

Richardson, A. D. (2023). PhenoCam: An evolving, open-source tool to study the temporal and spatial variability of ecosystem-scale phenology. In *Agricultural and Forest Meteorology* (Vol. 342). Elsevier B.V. https://doi.org/10.1016/j.agrformet.2023.109751

Schädel, C., Seyednasrollah, B., Hanson, P. J., Hufkens, K., Pearson, K. J., Warren, J. M., & Richardson, A. D. (2023). Using long-term data from a whole ecosystem warming experiment to identify best spring and autumn phenology models. *Plant-Environment Interactions*, *4*(4), 188–200. https://doi.org/10.1002/pei3.10118

Scheifinger, H. (2023). *PEP725, the European phenological database*. https://doi.org/10.5194/egusphere-egu23-8348

Schelhaas, M., & Clerkx, A. P. P. M. (2014). *State of the forests in the Netherlands, 2012-2013*. https://www.wur.nl/en/Publication-details.htm?publicationId=publication-way-343832353034

Smyth, T. A. G., Wilson, R., Rooney, P., & Yates, K. L. (2022). Extent, accuracy and repeatability of bare sand and vegetation cover in dunes mapped from aerial imagery is highly variable. *Aeolian Research*, *56*. https://doi.org/10.1016/j.aeolia.2022.100799

Song, G., Wang, J., Zhao, Y., Yang, D., Lee, C. K. F., Guo, Z., Detto, M., Alberton, B., Morellato, P., Nelson, B., & Wu, J. (2024). Scale matters: Spatial resolution impacts tropical leaf phenology characterized by multi-source satellite remote sensing with an ecological-constrained deep learning model. *Remote Sensing of Environment*, *304*. https://doi.org/10.1016/j.rse.2024.114027

Templ, B., Koch, E., Bolmgren, K., Ungersböck, M., Paul, A., Scheifinger, H., Rutishauser, T., Busto, M., Chmielewski, F. M., Hájková, L., Hodzić, S., Kaspar, F., Pietragalla, B., Romero-Fresneda, R., Tolvanen, A., Vučetič, V., Zimmermann, K., & Zust, A. (2018). Pan European Phenological database (PEP725): a single point of access for European data. *International Journal of Biometeorology*, *62*(6), 1109–1113. https://doi.org/10.1007/s00484-018-1512-8

Tian, F., Cai, Z., Jin, H., Hufkens, K., Scheifinger, H., Tagesson, T., Smets, B., Van Hoolst, R., Bonte, K., Ivits, E., Tong, X., Ardö, J., & Eklundh, L. (2021). Calibrating vegetation phenology from Sentinel-2 using eddy covariance, PhenoCam, and PEP725 networks across Europe. *Remote Sensing of Environment*, *260*. https://doi.org/10.1016/j.rse.2021.112456

Wang, J., Song, G., Liddell, M., Morellato, P., Lee, C. K. F., Yang, D., Alberton, B., Detto, M., Ma, X., Zhao, Y., Yeung, H. C. H., Zhang, H., Ng, M., Nelson, B. W., Huete, A., & Wu, J. (2023). An ecologically-constrained deep learning model for tropical leaf phenology monitoring using PlanetScope satellites. *Remote Sensing of Environment*, *286*. https://doi.org/10.1016/j.rse.2022.113429

Zhang, J., Qiu, X., Wu, Y., Zhu, Y., Cao, Q., Liu, X., & Cao, W. (2021). Combining texture, color, and vegetation indices from fixed-wing UAS imagery to estimate wheat growth parameters using multivariate regression methods. *Computers and Electronics in Agriculture*, *185*. https://doi.org/10.1016/j.compag.2021.106138

Zhu, Y., Chen, M., Gu, Q., Zhao, Y., Zhang, X., Sun, Q., Gu, X., & Zheng, K. (2022). Machine learning methods for efficient and automated in situ monitoring of peach flowering phenology. *Computers and Electronics in Agriculture*, *202*. https://doi.org/10.1016/j.compag.2022.107370