

Review article

A review of semantic segmentation methods and their application in apple disease detection



Masoumeh Keshavarzi ^a, ^{*}, Carl Mesarich ^a, Donald Bailey ^b, Martin Johnson ^c,
Gourab Sen Gupta ^d

^a School of Agriculture and Environment, Massey University, Palmerston North, 4442, New Zealand

^b Christchurch, New Zealand

^c School of Mathematical and Computational Sciences, Massey University, Albany, Auckland, 0632, New Zealand

^d School of Food Technology and Natural Sciences, Massey University, Palmerston North, 4442, New Zealand

ARTICLE INFO

MSC:
0000
1111

Keywords:

Semantic segmentation
Deep learning
Transformers
Apple disease detection

ABSTRACT

Semantic segmentation, with pixel-wise classification, enables the precise identification of different parts of plants, as well as the diseases that occur on them, in agricultural images. Apples, as one of the most important fruit crops worldwide, are susceptible to various diseases, causing decreased crop quality and increased crop loss. To prevent disease progression and ensure prompt treatment, semantic segmentation acts as an effective method in the context of apple disease detection. This review provides a comprehensive analysis of semantic segmentation methods applied in apple disease detection, ranging from traditional approaches to state-of-the-art techniques. By systematically examining the entire pipeline, from dataset preparation to the segmentation and evaluation stages, this work not only synthesises existing knowledge but also reviews applied solutions and highlights remaining research gaps to enhance segmentation performance. Additionally, it offers a forward-looking perspective by proposing future research directions. Overall, this review aims to advance plant disease detection through semantic segmentation, with a particular emphasis on apples.

Contents

1.	Introduction	2
2.	Methodology of literature selection	3
3.	Segmentation methods	4
3.1.	Traditional methods	4
3.1.1.	Threshold-based methods	4
3.1.2.	Colour-based methods	4
3.1.3.	Clustering-based methods	4
3.1.4.	Limitations	6
3.2.	CNN models for semantic segmentation	6
3.3.	Vision transformers for semantic segmentation	8
4.	Exploring challenges, recent solutions and gaps	9
4.1.	Dataset preparation	9
4.1.1.	Challenge : Limited available data	9
4.1.2.	Solutions	9
4.1.3.	Gaps	14
4.2.	Model development	15
4.2.1.	Challenges: Boundaries, small objects, imbalanced data	15
4.2.2.	Solutions	15
4.3.	Evaluation of the segmentation results	18
4.3.1.	Measurement metrics	18
4.3.2.	Ablation studies and visualisation	19

* Corresponding author.

E-mail address: M.keshavarzi@massey.ac.nz (M. Keshavarzi).

4.3.3. Gaps	19
5. Multispectral and hyperspectral imaging	19
6. Apple disease severity assessment	20
7. Costs of semantic segmentation	20
7.1. Data collection effort	20
7.2. Data preparation and processing	21
7.3. Computational costs	21
8. Conclusion and future direction	21
CRedit authorship contribution statement	22
Declaration of competing interest	22
Acknowledgement	22
Data availability	22
References	22

1. Introduction

Apples are a globally popular and significant fruit crop, but their quality and yield may decrease due to their susceptibility to several diseases (Tian et al., 2021). Some major diseases, such as Alternaria spot or blotch, primarily develop on leaves, resulting in extensive defoliation that indirectly affects fruit size and shape (Grove et al., 2003), while others, like scab, also lead to defoliation, but have a more direct impact on the fruit through the large-scale development of fruit lesions (Bowen et al., 2011). Since consumers value fruit based on its appearance and taste, these changes in quality can ultimately lead to a product that is unmarketable (Zanetti et al., 2020). Notably, certain apple diseases, such as scab, also demand substantial chemical intervention every season for their control (Bowen et al., 2011). The extensive or exclusive use of these pesticides, however, can negatively impact the health of the consumer and the environment (Bowen et al., 2011). To assess the severity and progression of apple diseases, and therefore determine the appropriate timing and usage of pesticide sprays needed for their control, as well as to mitigate yield and economic losses, the rapid and accurate detection and identification of apple diseases is essential (Chatzidimopoulos et al., 2020; Storey et al., 2022).

In the visual disease detection process, experts monitor plants for symptoms and then aim to match the symptoms with a particular disease to make a diagnosis. This process, however, is subjective and time-consuming, especially in orchards. Recently, machine vision-based systems have been used for plant disease detection, providing a faster and more accurate diagnosis to ensure timely control (Shahoveisi et al., 2023). Deep learning (DL)-based vision systems have demonstrated remarkable performance in image processing tasks in different domains (Kotwal et al., 2023). With their ability to automatically learn relevant features from data, convolutional neural networks (CNNs) have revolutionised image analysis in the agricultural domain by achieving state-of-the-art results. Different methods are applicable to plant disease identification and classification, including those associated with fruit (Zhang et al., 2022) and leaves (Yadav et al., 2022; Tang et al., 2023; Li et al., 2022b; Bansal et al., 2021; Liu et al., 2018; Nachtigall et al., 2020). Localisation methods usually define the area of disease through bounding boxes and are widely used in lesion detection on different parts of apple trees (Khan et al., 2022a; Di and Li, 2022; Hou et al., 2023; Zhang et al., 2023c). The same methods are commonly used in weed identification, given that weeds act to harm or prevent proper plant growth (Gallo et al., 2023). While classification and localisation methods assign labels to individual images or objects within them, there are certain key issues they cannot address. In particular:

- Determining the precise boundary of each leaf, fruit, or disease lesion.
- Quantifying the percentage of leaf surface affected by a disease.
- Providing detailed and pixel-level segmentation, to distinguish between a leaf, fruit, stem, or disease lesion.

Agricultural images often contain complex backgrounds and different objects with similar colour variations. Detecting fine-grained features from such images needs methods capable of separating out the various objects and regions into meaningful parts. Results have shown that separating targeted areas (leaves, fruits and lesions) from complex backgrounds can facilitate disease classification (Dubey and Jalal, 2012; Bracino et al., 2020) or improve the classification results (from 95.36% to 98.79% in Zhang et al. (2023a), from 80.6% to 84.3% in Yu et al. (2019), from 93.4% to 96.9% in Khan et al. (2018), from 97.50% to 98.63% in Hasan et al. (2022b), from 42% to 46% in Mzoughi and Yahiaoui (2023)). Semantic segmentation has been used to assign a label to each pixel, indicating the object or class to which it belongs (Abdalla et al., 2019; Gonçalves et al., 2021; Jamadar and Sharma, 2023; Sodjinou et al., 2022). This allows further processing, such as an assessment of disease severity in the crop (e.g. as shown for rice Chen et al., 2021, cucumber Wang et al., 2021, and grape Liu et al., 2022b). In the context of apple diseases, there is a need for meaningful parts (any type of lesion and pathogen, or the whole region of a leaf, apple, or stem) to be precisely detected. Fig. 1 shows the difference in apple disease detection through different methods. Segmentation is also critical in detecting apple diseases, acting as a foundational step for further analysis. By defining the exact location of the diseased area, farmers can promptly implement targeted measures that prevent the infection from spreading further (Rehman et al., 2021).

Some recent studies in apple disease segmentation are highlighted by Bonkra et al. (2024). This bibliometric analysis showed the importance and increasing trends of segmentation and machine learning techniques in apple disease detection. However, the same article did not exclusively review segmentation techniques, challenges and solutions for apple disease detection. In another review by Barbole et al. (2022), the primary focus was on apple fruit segmentation to facilitate harvesting, while in a review by Lu and Lu (2017) only the detection of post-harvest apple fruit defects through various imaging technologies was a focus; however, in the latter review, the detection of leaf and preharvest apple lesions was not covered. Multispectral systems for tree disease detection were reviewed by Fang et al. (2023). Rayhana et al. (2023) discussed hyperspectral imaging specifically for plant disease detection. Eh Teet and Hashim (2023) also provided an review of optical imaging techniques and their application in detection of diseases in fruits including apples and vegetables. However, the focus in these study was imaging techniques rather than segmentation. Finally, in a review by Bonkra et al. (2023), a limited aspect of segmentation in detecting apple diseases was addressed. However, this did not encompass the entire spectrum of segmentation models used in apple disease detection, nor did it cover the challenges encountered at various stages of the process and applied solutions.

Main contributions of the current review :

- Semantic segmentation methods in the context of apple disease detection across both pre and post-harvest stages, and covering both leaves and fruits, are surveyed.
- Application of semantic segmentation in apple disease detection are provided.

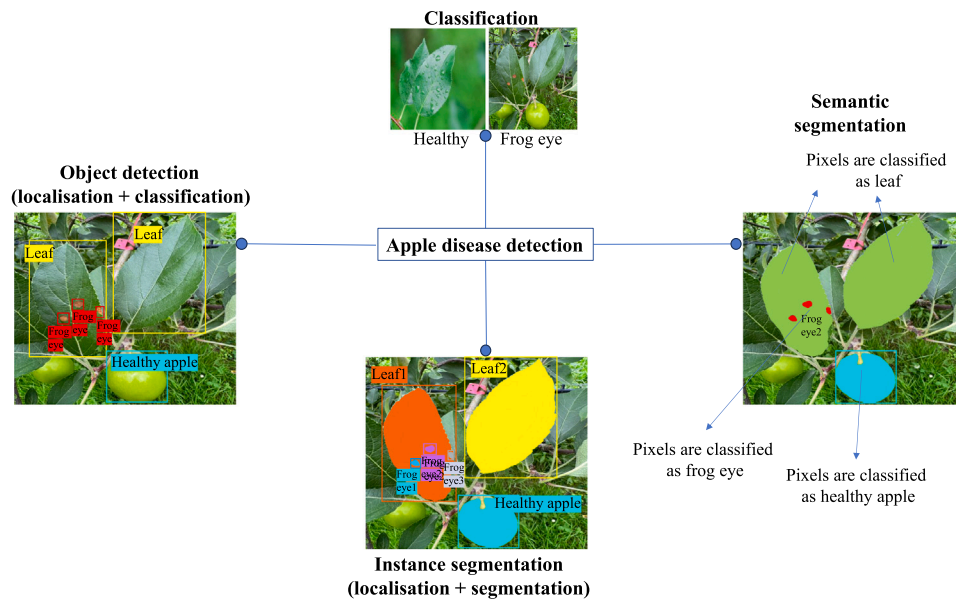


Fig. 1. Different DL-based methods in apple disease detection.

- Key gaps in datasets, annotations, model architectures, and evaluation practices are highlighted and some techniques employed in recent studies to fill these gaps are outlined.
- The cost associated with semantic segmentation from a digitisation footprint perspective is discussed.
- Future research opportunities and possible directions for progress in the field are outlined and discussed.

The objective of this review is to examine a wide range of semantic segmentation methods applied to apple disease detection and to provide a comprehensive overview that can assist researchers in selecting the most appropriate approaches for their specific tasks. Through this comprehensive investigation, we contribute to the selection of more accurate and efficient methods for the advancement of apple disease detection technologies in disease severity assessment in the agricultural sector.

The remaining sections of this review are organised as follows. Section 2 of this paper includes an explanation of how literature was retrieved and reviewed in a systematic manner. In Section 3, the methods of segmentation are comprehensively reviewed. This allowed us to highlight gaps in datasets, annotation methods, model architectures, and evaluation methods mostly associated with the semantic segmentation of apple diseases in Section 4. Apple disease detection through multi and hyper-spectral imaging, and apple disease severity assessment are covered in Sections 5 and 6 respectively. The cost of semantic segmentation is discussed in Section 7. Finally, in Section 8, a conclusion and future directions are provided.

2. Methodology of literature selection

To identify the articles relevant to computer vision based apple disease segmentation methods, a systematic literature review was undertaken. Papers related to apple disease detection, semantic segmentation methods, and apple disease severity assessment were retrieved. For this purpose, the major academic databases used were ScienceDirect, Google Scholar, IEEE Xplore, SpringerLink, and the ACM Digital Library. The search covered publications from 2002–2025, using keywords listed in Table 1.

Of 480 publications, 216 papers were chosen. The selection criteria focused on the following:

- studies on apple disease detection employing segmentation methods associated with

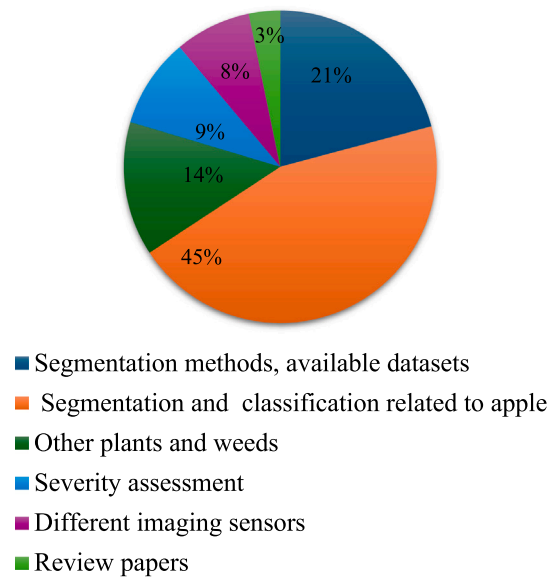


Fig. 2. Distribution of the proportions of the available publications used in this review.

- either in-field or laboratory settings,
- pre- and post-harvest stages
- studies involving other plant types were included if
 - they contributed to a broader understanding of the segmentation approaches relevant to this domain.

To enhance the readability of the review, the selected papers were categorised into the different categories as shown in Fig. 2. After conducting a thorough review of the selected publications, the findings are presented in the subsequent sections of the review.

Table 1
Overview of search criteria used for literature retrieval.

Category	Search terms ^a
Detection	“apple leaf” OR “fruit” AND (“post-harvest defects” OR “disease detection”)
Methods	“apple” OR “fruit” AND (“semantic segmentation” OR “hyperspectral images”)
Application	“agriculture technology” OR “deep learning in agriculture” OR “apple” OR “fruit” AND (“disease severity” OR “severity assessment tools” OR “disease progress”)

^a Some search terms were used in combination with category. e.g., “apple disease detection”. Some terms were used interdependently. e.g. Application category with Methods category.

3. Segmentation methods

There are various techniques available for the segmentation of indoor and outdoor images of apple fruits and leaves. Segmentation aims to partition these images into distinct regions, facilitating various applications such as fruit harvesting (Jia et al., 2020; Selvaraj and Farzan, 2024; Kang and Chen, 2019), fruit counting, yield and fruit size estimation (Bargoti and Underwood, 2017; Häni et al., 2019; Gené-Mola et al., 2023), pruning (Tong et al., 2023), spraying (Liu et al., 2024b; Gao et al., 2020; Xue et al., 2023), as well as weed (Milioto et al., 2018; Sodjinou et al., 2022; Zou et al., 2021; Gupta et al., 2023; Moazzam et al., 2023; Zou et al., 2023) and disease identification (Yu et al., 2019; Chen et al., 2021; Li et al., 2022a; Sun et al., 2024; Wang et al., 2021; Liu et al., 2024a; Shoab et al., 2022; Mzoughi and Yahiaoui, 2023; Khan et al., 2022b; Xu et al., 2022; Zhang and Zhang, 2023).

3.1. Traditional methods

When using traditional machine vision techniques, segmentation plays a crucial role in crop disease classification (Fig. 3). Prior to subsequently extracting features and identifying diseases, segmentation can be employed for extracting diseased areas within the image.

3.1.1. Threshold-based methods

Traditional threshold-based segmentation methods are straightforward and widely used to generate a binary segmented mask. The fundamental objective is to distinguish between two classes in the image, typically separating foreground (e.g., plant leaf or disease) from background, using a threshold value. This method’s binary nature does not provide precise labels for individual parts or pixels and is often insufficient for more complex segmentation tasks that involve multiple classes or disease types. Identifying an appropriate threshold value that can effectively cover variations in agricultural images is a complex task. One significant issue arises from the diversity of vegetation images, which can exhibit considerable colour variations. Most research has attempted to find a threshold value that is best tailored to particular datasets, leading to poor generalisation when working on new samples from different environments. Some thresholding methods, such as Otsu (1979), which are commonly used to separate leaves from simple and uniform backgrounds, lead to missegmentation by considering lesions as the background. Upadhyay and Kumar (2022) revealed that Otsu thresholding is effective for separating leaf from simple background and increases the classification accuracy to 97.7%. However, segmentation of the lesion pixels was affected by the background with colour similarity. In another study, working on defect detection in apple fruits at post-harvest, Fan et al. (2020) reported that non-uniform illumination across different samples may lead to darker parts of the apple being classified as the defect area. Qualitative results revealed that the thresholding method could not distinguish true defects from stem/calyx ends because of the similar intensity values.

3.1.2. Colour-based methods

Colour-based methods segment images through colour intensity. Each pixel is segmented into the same colour area based on its surrounding pixels. These methods are sensitive to light changes and colour variations, making them less robust for application in apple disease segmentation and detection. While there are some successes associated with colour-based methods for object segmentation in agricultural images, they can usually identify objects having significant and different colours, known in advance, from surroundings such as red apples. In Fan et al. (2021), testing on 180 apple images showed an average accuracy rate of 99.26%. Still, some false positive results were observed in separating leaves from the background within a similar colour range (see Fig. 4). Hu et al. (2023) extracted colour features from the hue, saturation, and value (HSV) colour space in tea leaves as input for a support vector machine (SVM) classifier with an increase in mean of intersection over union (MIOU) to 79%. However, highly illuminated parts of the leaves were segmented wrongly as disease. Some preprocessing steps such as contrast limited adaptive histogram equalisation (CLAHE) could alleviate this by normalising the pixels’ intensity.

GrabCut (Rother et al., 2004) is a colour-based and interactive method widely used in agricultural image segmentation. In this method, images are segmented based on the colour and distance relationship between pixels to create meaningful regions or objects. Probability graphs are obtained for two Gaussian mixture models (GMM). An illustration of the constructed graph in the GrabCut method is shown in Fig. 5. As thicker foreground and background edges have a higher probability and cost of cutting, the cutting process is done at low-cost edges to leave a bigger segment (Wang et al., 2023).

Hasan et al. (2022a) utilised a combination of the GrabCut algorithm for apple leaf segmentation and a colour-based thresholding method for leaf lesion segmentation with an MIOU of 83%. In another study, Pal and Kumar (2023) proposed the AgriDet framework in which farmers can upload a image to get the disease severity. In this work, the images from Plantdoc (Singh et al., 2020) and Plantville (Mzoughi and Yahiaoui, 2023) datasets were considered with simple and complex background images, respectively. A multivariate GrabCut algorithm was used to segment apple leaves and lesions for classification of disease severities. While qualitatively, there was poor accuracy in segmenting the whole leaf from complex backgrounds, the framework showed a good final classification of apple scab and rust disease severity (95.84% and 85.75% respectively). Lian et al. (2024) added post-processing steps to refine the GrabCut segmentation result of apple leaves. A median filter was used to eliminate noise and a contour function was employed to define the biggest connected area within the image (a single leaf). Fig. 6 demonstrates the generated mask with both GrabCut and their modified C-GrabCut algorithm.

3.1.3. Clustering-based methods

Clustering-based techniques, especially the K-means approach, are employed for complex image segmentation. However, their performance depends heavily on the number of chosen clusters and their initial centroids. Results have shown that K-means clustering method usually leads to some errors in disease segmentation (Javidan et al., 2023; Larjani et al., 2019; Hu et al., 2023). Colour and clustering-based methods often define diseases based solely on the most dominant colour in the centre of the lesions, which provides a sub-optimal segmentation. This is not beneficial because apple diseases with a fragmented or spotty appearance, such as rust and frog eye, typically exhibit diverse colour patterns with variation from their centre to the surrounding regions. Typically, additional post-processing operations are required to refine segmentation results and capture the entire lesion. Abuhayi and Mossa (2023) argued that customised segmentation with hybrid methods of thresholding and post-processing leads to more precise segmentation and higher classification accuracy (from 86.05% to 93.59%).

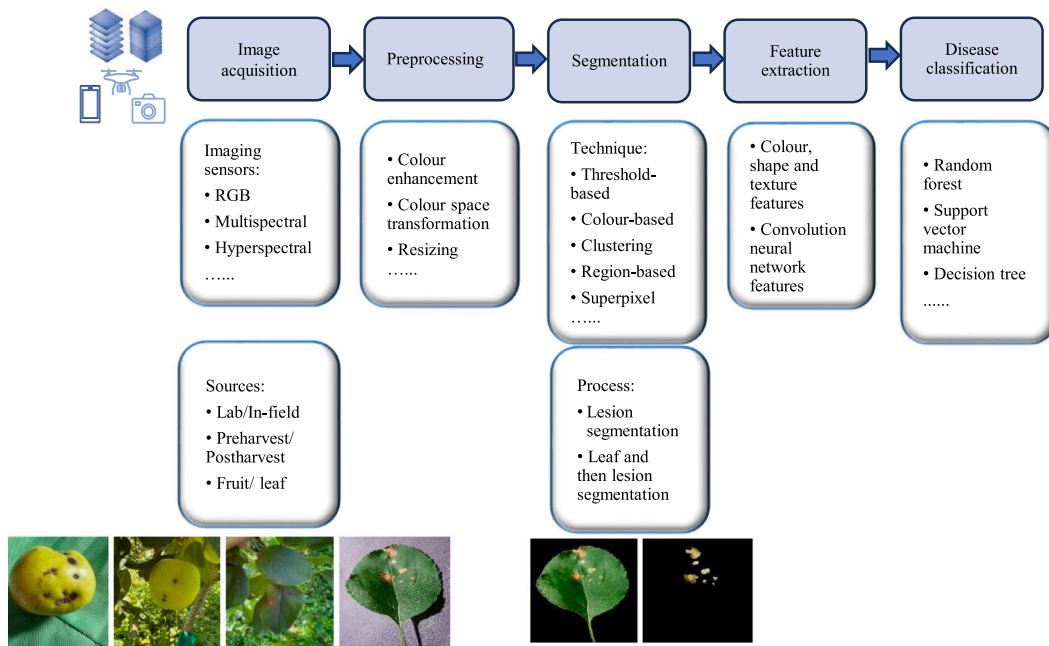


Fig. 3. The process of traditional computer vision method in apple disease classification and the role of segmentation before feature extraction.



Fig. 4. Colour-based segmentation of an object with distinct colour regions. Colour similarity in the background leads to over-segmentation of the apple. Source: Image from Fan et al. (2021).

Superpixel segmentation methods segment images into meaningful parts by aggregating similar pixels. For example, simple linear iterative clustering (SLIC) method (Achanta et al., 2012), as one of the most used superpixel algorithms, has adapted the K-means clustering algorithm to efficiently group pixels into superpixels based on their colour similarity (in the LAB colour space) and spatial proximity. The SLIC algorithm requires the number of superpixels (K) to be defined as a hyper-parameter to cluster the pixels (N). It is important to choose an appropriate number of superpixels to provide a proper segmentation covering the maximum similar pixels in a superpixel. There is no optimal technique except using empirical guidelines or trial and error to determine the suitable number of clusters and superpixels. Fig. 7 (Zhang et al., 2019) shows good SLIC performance on a single leaf in which similar diseased pixels are clustered together. Nevertheless, some superpixels contain both healthy and unhealthy pixels. In a comparative study, to show the effectiveness of superpixels, Kim (2021) fed derived superpixels from SLIC into a CNN model to classify scab, black rot and cedar rust of apple leaves. Results showed that superpixels could be comparable with the full images in disease classification. Some studies leveraged a combination of clustering methods such as SLIC and K-means to perform the segmentation in tomato (Khan and Narvekar,

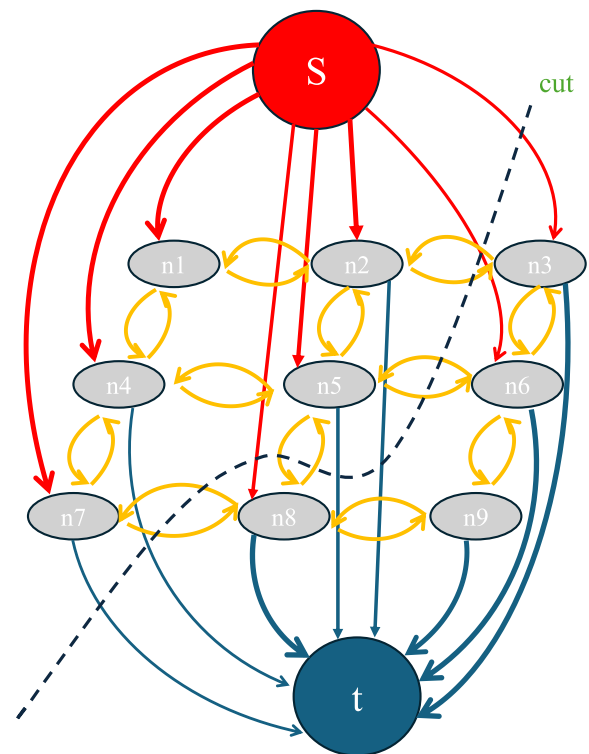


Fig. 5. GrabCut diagram showing the relationship between pixels and the probability of these pixels belonging to the foreground and background segment. The s and t nodes show the foreground and background nodes, respectively. The thickness of the edges represents the weight of the edge. Source: Image from Wang et al. (2023).

2022), apple and cucumber leaves (Zhang et al., 2018). Not considering the background as a separate cluster resulted in over-segmentation by considering background objects as lesions. Some other similar parts of the image (e.g. veins, shadows or background) are more likely to be considered as lesions.

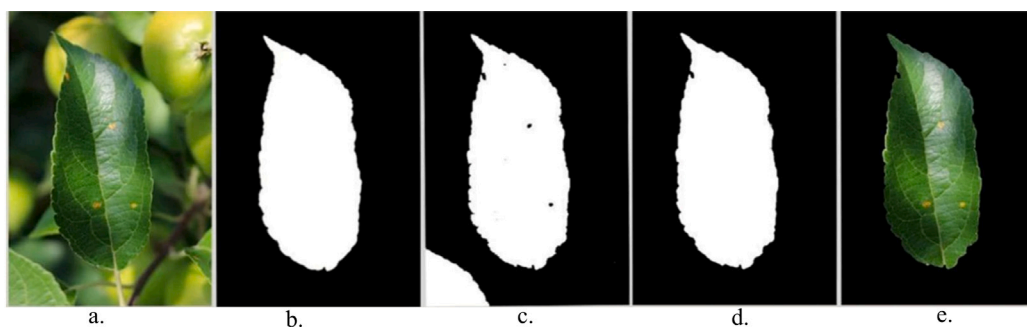


Fig. 6. Mask generation of GrabCut algorithm and the modified GrabCut. a. shows the original image, b. is the ground truth mask, c. illustrates the mask generated with GrabCut, d. represents the C-GrabCut mask and e. is the final segmented image.

Source: Image from Lian et al. (2024).

Region-based segmentation methods start with initial seeds and gradually put neighbouring pixels into each seed category based on similarity. This similarity usually is based on the intensity of the pixels. Region growing algorithm is one of the classical region-based methods that starts with one or more seed points and grows regions around these seeds by incorporating neighbouring pixels that satisfy certain criteria (Jourlin et al., 2013). This method provided an acceptable segmentation of clear and mature lesions on apple leaves against a simple background (Chuanlei et al., 2017). Another method under the region-based segmentation category is the marker-watershed algorithm. In this method, initial points (markers) need to be defined either manually or automatically. After transforming images into grayscale, this method creates a height map that highlights disconnected parts such as sudden intensity changes at object boundaries and dampens continuous regions including parts of the image with similar intensity. The algorithm then starts flooding from the initial markers (Fan et al., 2019). This algorithm also outperformed the Otsu and K-means methods in extracting tea diseased leaves (MIOU = 83%) (Hu et al., 2023). The Marker-watershed method is effective when object and background intensities or gradients differ clearly such as detecting the defect area of post-harvest apples showing better results than Otsu and K-means segmentation methods (Yogesh et al., 2017). As a practical application, this technique was used as a postprocessing technique to distinguish between connected segmented apples. The RGB images were captured by an unmanned ground vehicle in the apple orchard for fruit counting and yield estimation. Comparing fruit counts with the number of harvested apples, the F1-score was 86.1% (Bargoti and Underwood, 2017).

3.1.4. Limitations

In terms of apples, young fruit with green colour are not easily distinguishable from leaves using colour-based segmentation methods. Furthermore, outdoor samples tend to also have differing backgrounds, e.g. soil, which might be predicted as a lesion, or dust and raindrops on the surface of a healthy leaf or fruit, which can be incorrectly identified as a disease. It should also be noted that highly illuminated and saturated pixels, due to their similarity in colour to some apple leaf diseases, such as those associated with a mosaic virus, or immature rust and scab infections, might be misidentified as these diseases (Fig. 8 provides a detailed illustration).

While there have been advances in leaf and lesion segmentation for plant and apple disease, colour variations and illumination differences, complex backgrounds, and the need for precise pixel boundary identification are the main challenges for traditional segmentation methods. In-field fruit, leaf and disease segmentation have challenges in region detection due to colour similarities between different objects. Most thresholding, clustering and colour-based methods usually fail to accurately identify the boundary between the healthy region with various shades of green colour and the diseased area. As shown in the visual results of a recent study, poor segmentation of in-field images

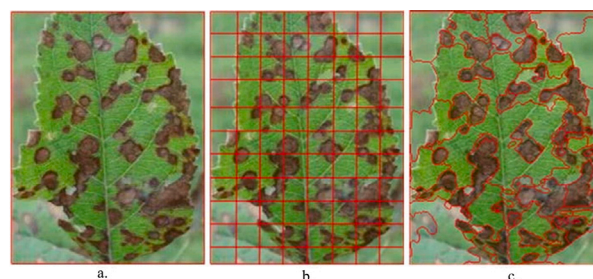


Fig. 7. Unhealthy leaf segmentation with SLIC algorithm. a. is the original image, b. shows the divided image with grids and c. is the final derived superpixels on the image. Source: Image from Zhang et al. (2019).

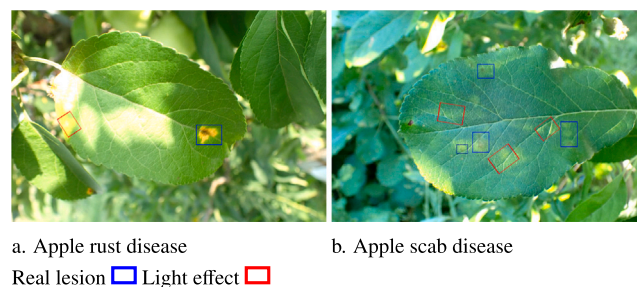


Fig. 8. Different illumination across a leaf surface can lead to lighting patterns that can be mistaken for immature rust and scab lesions. Source: Images from Thapa et al. (2021).

occurred, particularly in separating leaves from the background where colour variation existed within a single leaf (Parashar and Johri, 2024). Table 2 lists some studies where some traditional image processing approaches were used for apple fruit, leaf and disease segmentation, together with limitations.

3.2. CNN models for semantic segmentation

Recent research shows that the transition from traditional segmentation methods to DL-based semantic segmentation has resulted in substantial improvements in accuracy and generalisation (Sehar and Naseem, 2022; Kotwal et al., 2023; Khan et al., 2022b; Zhang and Zhang, 2023). In a CNN classification task, the input image goes through convolution layers followed by pooling layers to reduce spatial resolution and capture the most important features hierarchically. Fully connected layers are used towards the end of the network for classification. An image classification result is a one-dimensional vector with each component representing a different class. The largest component is used to assign a single label to the whole image. Long

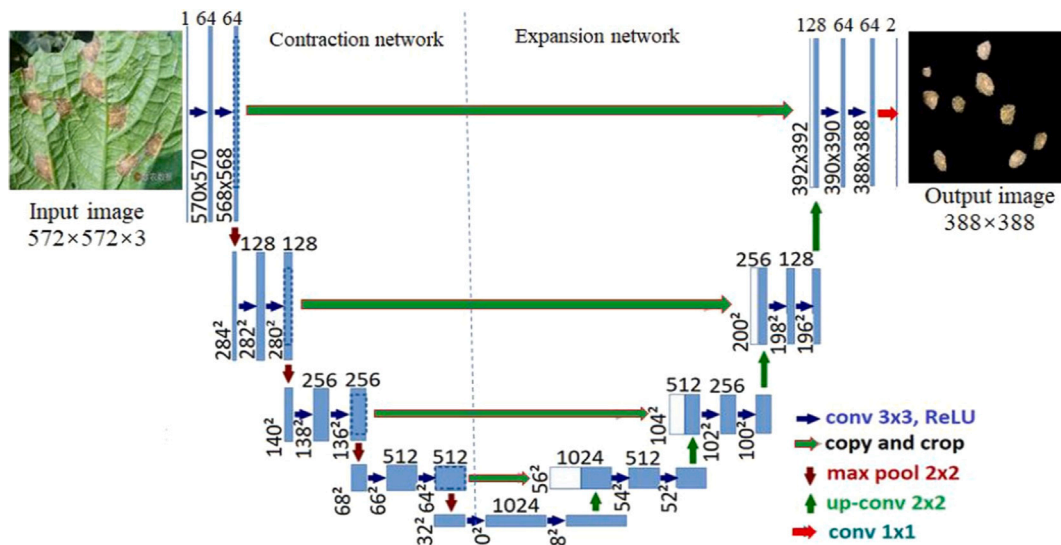


Fig. 9. U-Net contraction and expansion stages in plant disease segmentation. Source: Image from Zhang and Zhang (2023).

et al. (2015) extended this concept using fully convolutional networks (FCNs) to provide a pixel-wise classification. Instead of producing a single label for the entire image, FCNs produce a dense prediction map where each pixel is classified. The whole process of FCNs is based on encoder and decoder blocks. The encoder block first identifies and encodes significant features through convolution layers, and the decoder block then reconstructs the segmentation output through deconvolution layers. Recently, significant progress has been made in the field of semantic segmentation using DL, and many researchers use FCNs and meta-architecture models including U-Net (Ronneberger et al., 2015), SegNet (Badrinarayanan et al., 2017), different versions of DeepLab (V1, V2, V3, V3+) (Chen et al., 2015, 2018a, 2017, 2018b), or pyramid scene parsing network (PSPNet) (Zhao et al., 2017) in segmenting different parts of plants and disease lesions in both controlled environments and in-field datasets.

Incorporating spatial information

Considering leaves, lesions and fruits in an image, spatial information such as edges and location are mostly extracted at the lower levels of the network. While reconstructing images, the information from the initial encoder layers could be missed. This can result in low accuracy, especially at the edges. In the U-Net model, skip connections are deployed to transfer the output of layers (before applying the max-pooling operation) from the contraction network to corresponding layers in the expansion stages to preserve the spatial information and benefit high-level and low-level features. The skip connections help the network to protect fine-grained details and localise features. Fig. 9 shows a U-Net model that has resulted in disease lesion segmentation. This model has been used in the segmentation of apple leaf and diseases using RGB images (MIOU = 98.32% for leaf and 92.05% for disease spots Liu et al., 2022a), apple leaves using multispectral images (Das et al., 2023), as well as other plant structures and disease types (Divyanth et al., 2023; Wang et al., 2021; Rai and Pahuja, 2023). In a real field deployment, Uryasheva et al. (2022) developed a U-Net model with an EfficientNet backbone to segment multispectral images of healthy and infected apple trees. A three-camera sensor node was used to capture images from eight different apple tree varieties. From the segmented images, a desktop application was developed to generate detailed vegetation index maps for individual trees. During the blooming season, the model was tested and achieved an IOU of 72%.

Following the encoder–decoder architecture, SegNet uses max-pooling indices from the encoder to reconstruct the segmented

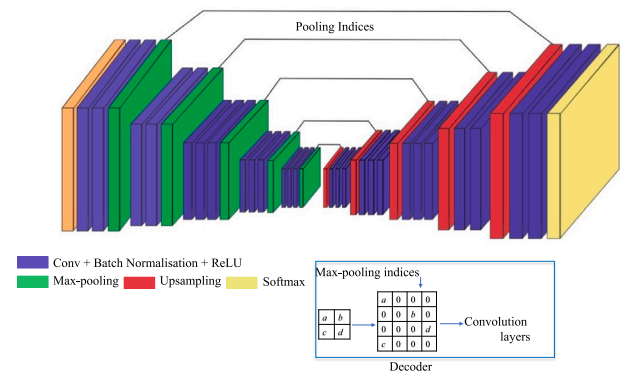


Fig. 10. Illustration of the SegNet model architecture and its decoder. The max-pooling indices are used for upsampling without learning the encoder’s feature map. The sparse constructed feature map then goes through the convolution layers. Source: Image adapted from Badrinarayanan et al. (2017).

output (Fig. 10 shows the structure). These max-pooling indices show the location of each maximum value in the pooling operation that makes this model more memory-efficient and faster compared to U-Net (with fewer parameters and reduced inference time for a single image Wang et al., 2021). In leaf and disease segmentation, the performance of SegNet is comparable to that of U-Nets, with slightly lower MIOUs (2.04% Divyanth et al., 2023, 0.8% Gonçalves et al., 2021, 0.35% Ngugi et al., 2020) or higher (4.15% Xu et al., 2020). The SegNet model was used on an intelligent spraying system in a pear orchard to determine fruit and tree segmentation with an accuracy of 83.79% by the Kim et al. (2020). To operate the nozzles, the authors separated each image captured from camera lengthwise into quarters and mapped these to the nozzles. The nozzle activated when fruit tree coverage in a zone exceeded 20%. Field Experiments showed effective reduction in used pesticide by targeting only fruit trees.

Preserving resolution and multi-scale feature extraction

Performing pooling operations or convolution layers with stride are strategies to capture information from larger scales. However, both techniques reduce the resolutions of output feature maps. Dilated convolution performs the convolution operation to capture information from larger scales without significantly increasing the number of parameters. It uses gaps or holes within the filter (Fig. 11), to have a

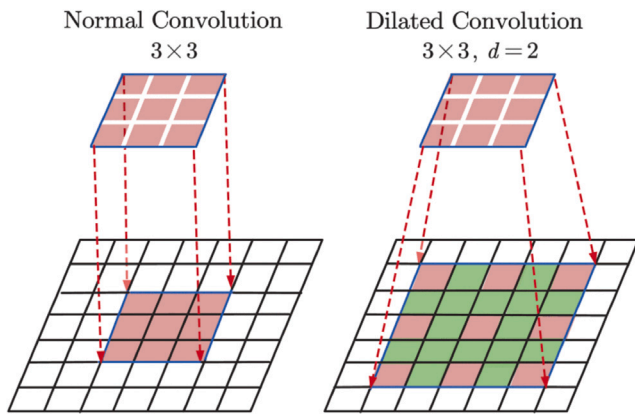


Fig. 11. The illustration of 2D dilated convolution with spatial size of 3×3 and dilation factor of 2. Both non-dilated and 2-dilated convolutions have the same number of parameters, whereas the dilated convolution has a larger receptive field. Source: Image from Du et al. (2020).

larger receptive field without increasing the kernel size. The dilation rate determines the spacing between the kernel's elements (pixels). This strategy to cover a larger scale for extracting features has been deployed in many apple disease detection models (Fu et al., 2022; Zhu et al., 2023a; Luo et al., 2021; Wang et al., 2024). DeepLabV1 and DeepLabV2 benefit from dilated convolution and fully connected conditional random field (CRF) in the segmentation models. CRF is a post-processing step after the initial segmentation to improve the results.

Atrous spatial pyramid pooling (ASPP) is a method that DeepLabV3 and DeepLabV3+ employ to capture multi-scale contextual information. This is obtained by using parallel atrous (dilated) convolutions with different dilation rates for pooling between different scales (Fig. 12). This technique can enhance segmentation performance by capturing multi-scale contextual information, which is particularly beneficial in agricultural images in which each object such as a leaf, lesion, fruit, etc. may appear at varying sizes (Zhu et al., 2023a; Liu et al., 2022a; Agarla et al., 2023). In practical development, Li et al. (2023a) introduced a Mobile application to enable the degree of apple leaf diseases (ring rot and rust) based on DeepLabV3+ model giving an MIOU of 97.26%. The deployed smartphone platform could predict the apple leaf diseases within 9 s. The DeepLabV3+ model showed an MIOU of 99.99% on predicting masks for decay zones in apples which is practical in tracking the dynamics of decay growth in apple fruits (Stasenko et al., 2021). Fig. 13 illustrates a DeepLabV3+ model employed for leaf and disease segmentation.

Global information

Focusing on the most relevant features globally, Zhao et al. (2017) proposed a four-level average pooling with different sizes: 1×1 , 2×2 , 3×3 , and 6×6 , to capture the features from different scales. Bilinear interpolation is used for upsampling to give the same size as the original feature map. Derived feature maps are then concatenated with the original one and is followed by the convolution layer to provide a pixel-level classification. This network usually delivers superior results in semantically separate objects that can help separate larger objects such as leaves in an agricultural image. According to Liu et al. (2022a), the PSPNet model had the highest performance in leaf segmentation (MIOU = 98.42%), while this model did not perform well in segmenting Alternaria spot lesions. Small lesions could not be detected and false positive predictions were made. Although PSPNet showed good accuracy in separating clear disease regions in images from the PlantVillage dataset, it showed poor performance in detecting fine-detailed information. Yellow tissue and light symptoms were not segmented properly at the ribs and margins of the leaves.

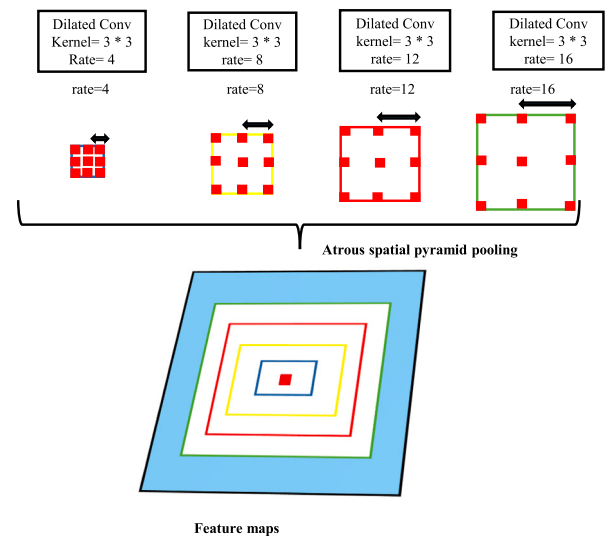


Fig. 12. Atrous spatial pyramid pooling (ASPP) to exploit multi-scale features by employing multiple parallel filters with different rates. Source: Image from Ma et al. (2019).

3.3. Vision transformers for semantic segmentation

The introduction of transformers has revolutionised natural language processing by introducing a self-attention mechanism that enables global and large-scale word dependencies to be detected. This led to the development of vision transformers (ViT) (Dosovitskiy et al., 2020) for classification and models such as SEgmentation TRansformer (SETR) (Zheng et al., 2021) and Segformer (Xie et al., 2021) for segmentation. CNNs utilise size-limited receptive fields (kernel sizes) to extract features, while Vision transformers divide the images into fixed-sized patches for processing. These patches are flattened and go through a trainable linear projection layer (a multilayer perceptron) to map into D dimensions in a process called patch embedding. A positional embedding is added to the patch embedding to provide the position information of patches. The SETR model shown in Fig. 14 is one of the initial transformer models for semantic segmentation tasks. Each transformation layer has a normalisation layer, a multi-head self-attention (MSA) which involves running k parallel self-attention operations, and a multilayer perceptron block. In a self-attention block, a weighted sum of embeddings is computed for each element, where the similarities between the components determine the weights. For each input token or image patch, a query q , a key k , and a value v vector are computed through linear transformations of the input embeddings. Then, a softmax function normalises the similarity scores to obtain attention weights. Finally, the decoder part uses convolution layers to reconstruct the image and yield a prediction. The decoding stage to generate the segmentation result is introduced in two different ways: progressive upsampling (Fig. 14.b) and multi-level feature aggregation (Fig. 14.c).

While SETR benefits from using ViT as the backbone with the patch size of 16×16 , the Segformer (Xie et al., 2021) model uses 4×4 patches, which results in a better dense prediction. Fig. 15 illustrates that the Segformer model utilises a hierarchical method in encoding patches and a lightweight multilayer perceptron in the decoder part. The patches go through the transformer blocks to obtain four multi-level features of coarse to fine-grained features with the size of $1/4$, $1/8$, $1/16$, $1/32$ of the image size. In each transformer layer, features from the attention module will pass through the Mix-FFN to have the location information with any input image size. The Mix-FFN is represented by

$$X_{out} = MLP(GELU(Conv_{3 \times 3}(MLP(X_{in})))) + X_{in} \quad (1)$$

where *MLP* is a multilayer perceptron, *GELU* is the activation function, and *Conv*_{3×3} is a 3 × 3 convolution layer. This prevents the model from being lost by using a fixed positional embedding in ViT and eliminates the need for additional interpolation operations in the case of larger images. Segformer and SETR models showed relatively good performance in segmentation of rice leaf diseases (Dai et al., 2024). Still, soils in the background with a similar colour were considered as diseases. The Segformer model and its different encoder variants showed satisfactory results in strawberry diseases (Elmessery et al., 2024). This model could extract the overall leaf features almost similar to DeepLabV3+ model due to the focus that transformer models have on global feature extraction (Lu et al., 2023). However, detailed evaluation revealed an inability in segmentation of small disease areas, diseases on broken leaves, serrations of the leaves (Lu et al., 2024), and diseased areas at the leaf tips (Lu et al., 2023).

Supervised DL and transformer models advance the plant leaf and disease segmentation, particularly in images with a complex background. Techniques such as concatenating the high- and low-level features, preserving feature resolution through dilated convolutions, and multi-scale feature extraction through pooling and dilated convolution layers with different rates, as well as patch processing, have all been employed to improve the segmentation results. However, challenges remain in capturing details, such as small disease lesions, diseases with similar regions to healthy parts, or those located at the margins. While the next section discusses the challenges, gaps, and solutions in details across different stages, Table 3 summarises the segmentation studies on apple fruit, leaf and their possible diseases using deep learning and transformer models, along with their limitations. Finally, Table 4 gives a summary of all reviewed techniques for semantic segmentation in this section.

4. Exploring challenges, recent solutions and gaps

The pipeline of DL-based semantic segmentation typically consists of three stages: (1) image acquisition and dataset preparation, (2) model development, and (3) result evaluation. While most of recent research in this area focuses on DL and transformer-based methods, this section highlights key challenges reported in the literature, the solutions proposed, and the remaining research gaps across the stages.

4.1. Dataset preparation

4.1.1. Challenge : Limited available data

In apple disease detection, leaves are often the first organ to show disease symptoms. Some researchers captured images from samples with simple backgrounds in laboratories, while others used in-field image samples against a complex background. In real scenarios, the models trained with lab-based images have poor generalisation to in-field disease detection (Ferentinos, 2018) (classification accuracy decreased by 66.26%).

As apple trees are deciduous, they only have leaves and fruit for a limited time during the year. This limits the ability to collect an image dataset that addresses all possible challenges including colour, illumination, and type of disease. Publicly available datasets for apple defect segmentation with multispectral images are scarce (Agarla et al., 2023) as these images usually have been used in classification tasks. While multispectral imaging requires specialised equipment and training, RGB imaging is more accessible and cost-effective. On the other hand, annotation of multispectral images is challenging which make them less efficient to be used in segmentation tasks. Consequently, numerous studies, especially those employing CNN and transformer-based segmentation models, have concentrated on disease detection using RGB images, which constitute the majority of the works reviewed in this study. Table 5 lists some datasets available for apple diseases. Several publicly available datasets contain RGB images annotated for classification and object detection, but few are annotated for segmentation. The existing labelled images are limited to a single masked leaf, disease, or the largest lesion. Furthermore, in some datasets, only images from leaves with simple backgrounds are provided.

4.1.2. Solutions

- To increase the generalisation of the model, both in-field and lab-based images should be considered. The low resolution of most images limits the precision of annotation, which in turn limits the accuracy of the results. Ahmad et al. (2023) demonstrated that the use of a combined dataset improved generalisation capability of deep learning models from 77.50% to 80.33%. Additional generalisation outcomes reported in related literature are also discussed in the aforementioned study.

- Supervised learning needs a large dataset in order to provide accurate predictions. One solution to high training time and the need for large datasets is transfer learning. Transfer learning leverages the features learnt by a different model pre-trained on a large dataset for a different (but ideally related dataset) task (Xu et al., 2022). Typically, the initial layers of the pre-trained model remain unchanged while the later layers are fine-tuned to extract features from task-specific data. When applied appropriately, transfer learning provides more accurate results in plant disease detection (Xu et al., 2020; Abdalla et al., 2019; Gonçalves et al., 2021; Barbedo, 2018; Xu et al., 2022) and apple disease segmentation (Li et al., 2023a; Lian et al., 2024; Mahmud et al., 2023; Rehman et al., 2021; Yu et al., 2019; Zhang et al., 2023b) than simply training with limited data. Transfer learning in semantic segmentation uses a pre-trained network, such as VGG-NET, AlexNet (Krizhevsky et al., 2017), or ResNet (He et al., 2016) in the encoder stage to accelerate the process of learning features, while the decoder part transforms distinctive features learned by the neural network from the input data.

- With only a small amount of related data, the chances of learning the most relevant features is reduced. A dataset can be augmented by creating new images from existing ones which can increase sample diversity, improve training and model accuracy, and prevent overfitting (Yang et al., 2023; Rai and Pahuja, 2023; Zhang and Zhang, 2023). Augmentation techniques include geometric transformations such as random horizontal and vertical flipping, rotations, resizing, colour changing, and noise addition (Gonçalves et al., 2021; Mzoughi and Yahiaoui, 2023; Jamadar and Sharma, 2023; Divyanth et al., 2023; Chen et al., 2021; Sheng et al., 2023; Feng et al., 2023; Storey et al., 2022; Zhang et al., 2023b; Lu et al., 2024). However, resizing and using images with a lower resolution can decrease the ability to detect small lesions (Tassis et al., 2021).

- Image augmentation through synthesised data generation models such as generative adversarial networks (GANs) is another solution for data scarcity which has improved apple disease segmentation and classification (Douarre et al., 2019; Tian et al., 2021, 2019). Nevertheless, background pixels may alter the shape and colour of the disease lesions and leaves in generating images. Chen et al. (2023) introduced an improved network to alleviate this problem, tested on the PlantVillage dataset. On the other side, outdoor images usually have more complexity in augmentation through GANs as results show that details are not preserved properly in some generated apple fruit images (Tian et al., 2021). Furthermore, to effectively generate data that closely resembles the training data, GANs require a larger number of training images. This reliance on large training image samples could be a drawback of the GAN techniques, as it contradicts the purpose of using data augmentation techniques, which primarily aim to address the scarcity of data. From this point of view, data augmentation could put a computational burden on the whole segmentation and disease detection process.

- Yang et al. (2022) leveraged few-shot learning to segment 10 types of diseases including apple leaf diseases. They proposed a method for generating prototypes from supported images using feature extraction techniques to segment new samples. A sub-set annotated dataset from the publicly available dataset (PlantVillage) was prepared. While they yielded a good result with a few samples, there were some missing results associated with small lesions. Few-shot segmentation of apple diseases on leaves was proposed in Bedi et al. (2024). In learning,

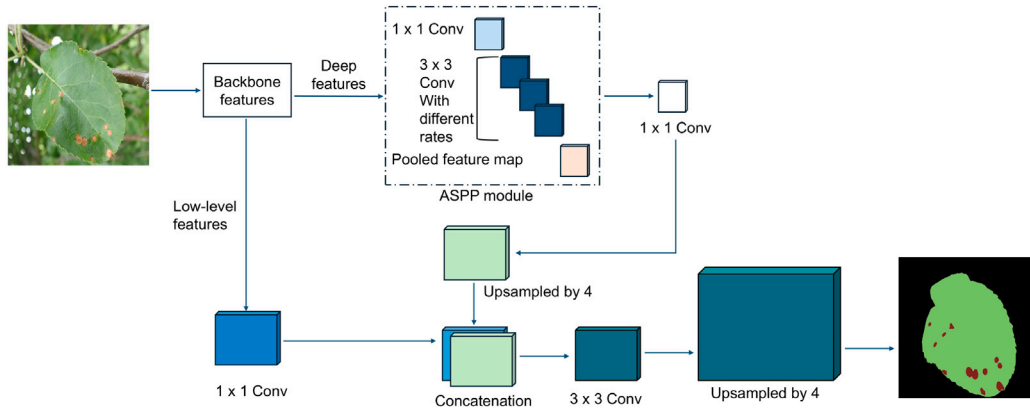


Fig. 13. Schematic diagram of the DeepLabV3+ model for leaf and disease segmentation.

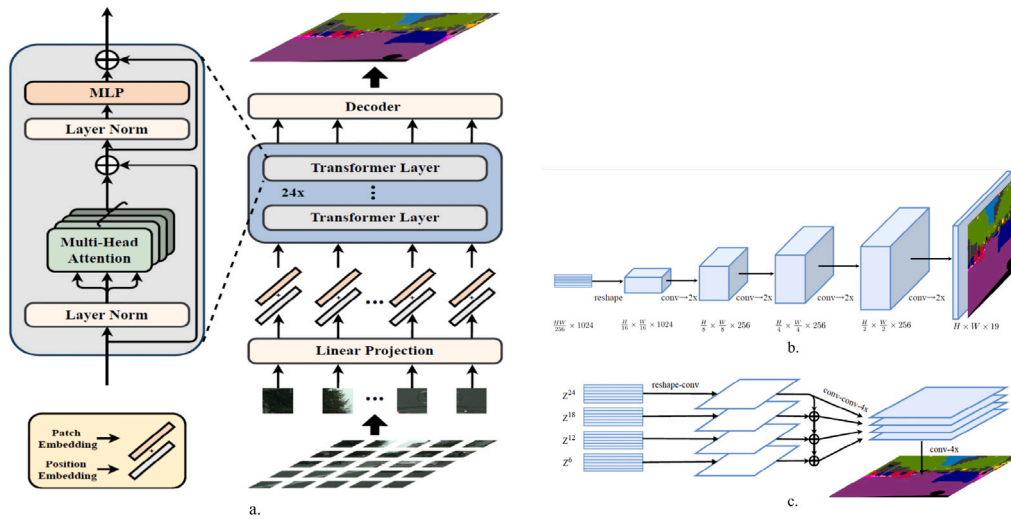


Fig. 14. Architecture of the Segmentation Transformer (SETR). a. Image patches are linearly embedded with position embeddings added, and then the resulting sequence of vectors is fed to standard transformer layers. b. Progressive upsampling through convolutions. c. Multi-level feature aggregation in which the transformer layers' outputs are reshaped and aggregated to form the prediction.

Source: Image from Zheng et al. (2021).

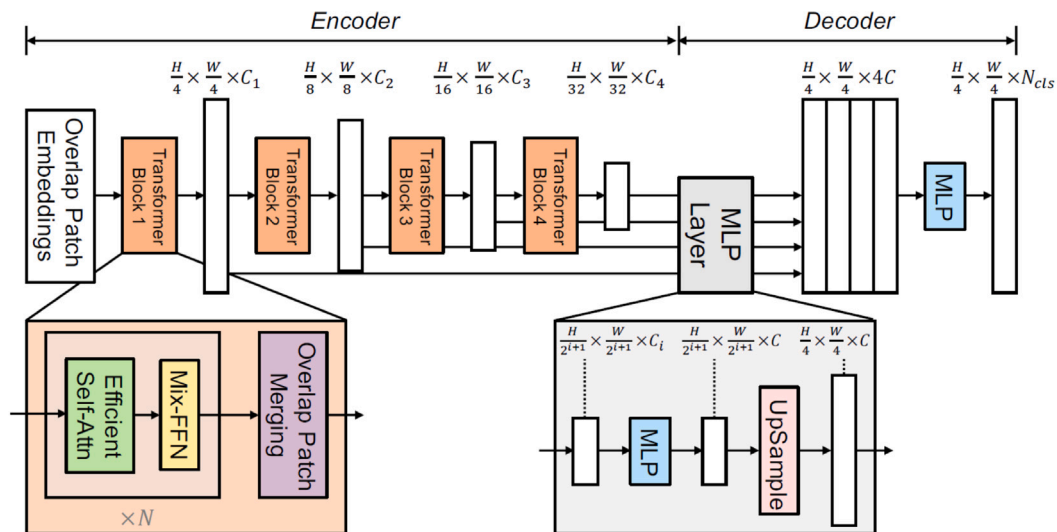


Fig. 15. Segformer architecture. Fine-grained and coarse features are extracted in a hierarchical encoder and a lightweight MLP fuses the multi-level features in the decoder to generate the mask.

Source: Image from Xie et al. (2021).

Table 2
Traditional segmentation methods in apple disease detection.

Ref	Dataset	Diseases	Methods	Results	Limitations
Dubey and Jalal (2012)	Own indoor images	Scab, rot, and blotch on fruits	K-means clustering for fruit segmentation	–	<ul style="list-style-type: none"> Needs optimal cluster number Not applicable in samples with different colour variations in apples
Bracino et al. (2020)	Plantvillage (Hughes et al., 2015)	Scab, rot, cedar apple rust on leaves	GraphCut to remove the leaf from the background	–	<ul style="list-style-type: none"> Not fully automatic Challenges with complex images and colour variations
Hasan et al. (2022b)	Plantvillage	Scab, black rot, and cedar apple rust on leaves	<ul style="list-style-type: none"> Utilised L*a*b* colour space for segmentation Extracted average colour markers in a*b* space Nearest neighbour method for segmentation of healthy, infected, and background regions 	–	<ul style="list-style-type: none"> Not applicable in complex background due to the colour similarity of classes (young disease lesions, background and healthy parts) Hard to define colour markers for all classes
Khan et al. (2018)	<ul style="list-style-type: none"> Plantvillage CASC IFW (Singh et al., 2011) 	<ul style="list-style-type: none"> Scab and apple rot on leaves Lesion on fruit with internal feeding worm 	A hybrid method comprises contrast enhancement, colour, texture, and correlation coefficient feature extraction to segment apple diseases	Avg-acc, 90.46% over 30 samples	<ul style="list-style-type: none"> Inaccurate segmentation of small lesions and multiple diseases Targeting only the largest lesion Requires post-processing step to ensure complete coverage of diseased area
Fan et al. (2020)	Own indoor images	Defects on apple fruits	Otsu's threshold method to remove the background	–	Misclassification between true defect and stem/calyx with a similar intensity value
Zhang et al. (2023a)	Own in-field images	apple with ring rot and Anthracnose	<ul style="list-style-type: none"> A pre-selection box for GrabCut segmentation is derived from the PSO algorithm K-means method to segment the apples from the background 	–	<ul style="list-style-type: none"> The bounding box limited the fruit selection when occluded with leaves Requires further processing to segment the apples
Hasan et al. (2022a)	In-field images from Son (2021)-	Marssonina blotch and Alternaria spot on leaves	Semantic segmentation using leaf separation with GrabCut segmentation and lesion segmentation via colour-based and thresholding methods	IOU 90%	<ul style="list-style-type: none"> Needs user action to segment the leaf Limited colours were considered in lesion segmentation Inaccurate border leaf and disease segmentation Detecting yellow leaf veins as lesions
Singh and Gupta (2018)	Own indoor and outdoor images	Marssonina coronaria and scab on leaves	<ul style="list-style-type: none"> Preprocessing: Dynamic histogram equalisation and image brightness normalisation Lesion segmentation: Green colour removal+K-means clustering 	–	Under-segmentation of leaf lesions with both simple and complex background
Zhang et al. (2018)	Own indoor images	Alternaria leaf spot, mosaic and rust on leaves	Combination of SLIC and K-means clustering	–	<ul style="list-style-type: none"> Requires the appropriate number of superpixels Lack testing the method on leaves with multiple diseases
Chuanlei et al. (2017)	Plantvillage	Powdery mildew, mosaic and cedar rust on leaves	<ul style="list-style-type: none"> Leaf segmentation with histogram method Lesion segmentation with region-based algorithm 	–	<ul style="list-style-type: none"> Needs to set an optimal threshold Under-segmentation in small and young rust lesion
Tian et al. (2024)	Own NIR images from harvested apples	Bruises on apples	Preprocessing to replace the calyx and stem with solid background + Adaptive threshold segmentation	F1-score 94.70%	Lose the chance of bruise detection around the stem and calyx
Parashar and Johri (2024)	Plant pathology 2020 (Thapa et al., 2020)	Healthy leaves, rust, scab, and multiple diseases on leaves	Canny and watershed algorithm to segment the leaves from background	–	Poor segmentation, especially in colour variation on a leaf

both indoor and outdoor images were considered. Although few-shot semantic segmentation shows promise in addressing data scarcity, the model is dependent on the few training samples with limited variation in light and background. The models' generalisation ability to segment diseases in test images with varying light and background conditions remains uncertain.

- Data annotation requires expert supervision, which makes the annotation of a large number of images a difficult and time-consuming task. It is usually a manual task, through annotation tools such as VGG Image Annotator (VIA), Labelme, or MATLAB Image Labeler, which can be prone to human error (Su et al., 2021; Khan et al., 2022b). Although

annotating a whole leaf or fruit needs accurate boundary selection, it is usually less complex than disease annotation. Apple disease annotation for segmentation tasks faces several challenges. Certain diseases such as rust and scab on leaves alter the surrounding colour area, raising the question: should subtle changes be considered as lesions? Therefore, there is a need for accurate and precise pre-defined disease recognition and definition before annotation.

Based on the data preparation process and various strategies employed to tackle challenges at this stage, some identified gaps are outlined as follows:

Table 3
CNN and transformers based methods in apple fruit, leaf and disease segmentation.

Ref	Dataset	Diseases	Methods	Results	Limitations
Liu et al. (2024a)	Plant pathology 2020 (Thapa et al., 2020)	Frog eye on leaves	<ul style="list-style-type: none"> Two-stage segmentation ResNet50+improved PSPNet with fewer pooling layers and using deformable convolution layers for leaf segmentation. ResNet50+dilated U-Net for lesion segmentation 	MIOU 91.27%, MPA 94.32%, precision 94.81%	<ul style="list-style-type: none"> A single leaf and disease type was considered The Mosaic image augmentation methods does not reflect some challenges of in-field images (e.g. high or low illumination, and brightness)
Mzoughi and Yahiaoui (2023)	Plantvillage	Scab, black rot, cedar apple rust on leaves	ResNet+PSPNet for symptom segmentation	IOU 61.44%	Mis-segmentation in lesions at ribes and leaf margins
Mahmud et al. (2023)	Own in-field images	Leaf fire blight	ResNet+MaskR-CCN	F1-score 91.96%	Under-segmentation by missing infected parts
Li et al. (2023a)	Own in-field images	Rust and ring rot on leaves	Improved Xception+DeepLabV3+	MPA 97.26%, MIOU 83.85%	<ul style="list-style-type: none"> Mis-segmentation of the leaf regions False segmentation in occluded leaves Only a single leaf was considered
Lu et al. (2024)	Own	Indoor apple leaf rust, outdoor leaves with Alternaria blotch and grey spot	<ul style="list-style-type: none"> EAIS-Former inspired by transformers & CNNs Dual scale overlap patch embeddings Replaced self-attention of transformer with ultra large scale transformer block Dual stage upsampling 	MIOU 99.04% for rust, 99.42% for Alternaria, 98.59% for grey spot	<ul style="list-style-type: none"> Lacks differentiation between disease labels Not considering multiple leaves and diseases Lacks testing the generalisation ability of the model
Sheng et al. (2023)	Own in-field images	Healthy green apples	<ul style="list-style-type: none"> Multiscale and global feature extraction Incorporating the edge information derived from pre-processing 	MIOU 90%	Poor segmentation in small apples
Wang and He (2022)	Own in-field images	Healthy apples	<ul style="list-style-type: none"> Image preprocessing to make the edges sharper and clearer Region Proposal Network to capture the region features Integrating deformable convolution and the transformer attention into Mask-RCCN 	F1-score 96.4%, MPA 91%	False segmentation in green apples and occluded with branches or leaves
Storey et al. (2022)	Plant pathology 2020	Leaf apple rust	ResNet50+Mask-RCNN	IOU 75%	<ul style="list-style-type: none"> Small number of training images Poor segmentation on the edges especially in blurry leaves Relative high rate in false positive by detecting healthy leaves as diseased
Son (2021)	Own in-field images	Marssonina blotch and Alternaria on leaves	FCN to leaf segmentation before the object detection module for leaf disease identification	–	<ul style="list-style-type: none"> FCN produced feature maps with over-segmentation and poor leaf boundaries Lacks differentiation between disease labels
Lu et al. (2023)	Own in-field images	Alternaria blotch and grey spot on leaves	A lightweight network (MixSeg) inspired by transformers and CNN to extract global and local features	Alternaria blotch IOU leaf 98.22% IOU disease 87.40% grey spot IOU leaf 98.09% IOU disease 86.20%	<ul style="list-style-type: none"> Lacks differentiation between diseases in disease segmentation Lacks segmentation for multiple diseases on leaves and segmentation for multiple leaves Lacks generalisation test
Zhu et al. (2023a)	Own	Alternaria, brown spot, grey spot, and rust on leaves (indoor and outdoor images)	Two-stage DeepLabv3+ with adaptive loss, adjusted dilation rates, and integrated attention mechanisms	IOU for leaf 98.70% IoU for disease 86.56%	Lacks multiple diseases on leaves and segmentation of multiple leaves
Zhu et al. (2023b)	MinneApple (Häni et al., 2020), AppleA (Dias et al., 2018)	Healthy apples, apple blossoms	<ul style="list-style-type: none"> Transformer-based model (PD-SegNet) Improved decoder with Dynamic Kernel Head and Complex Points Head to enhance the boundaries 	IOU 81.45%	Over-segmentation by considering green leaves as apples or flowers

(continued on next page)

Table 3 (continued).

Ref	Dataset	Diseases	Methods	Results	Limitations
Bedi et al. (2024)	ATLDS (Feng and Chao, 2022) (Indoor and outdoor images)	Alternaria, brown, grey, and rust spots on leaves	PDSE-Lite- A lightweight model for few-shot segmentation, used learned features from an autoencoder convolutional network	IOU 94.54%	<ul style="list-style-type: none"> Lacks generalisation assessment Poor segmentation in leaf boundaries Lacks visualisation in ablation study Lacks multiple diseases on leaves and segmentation of multiple leaves
Yang et al. (2024)	Own indoor and outdoor images	Alternaria, brown spot, grey spot, mosaic and rust on leaves	A U-Net based model with SegFormer blocks in the backbone and attention mechanisms in bottleneck and skip connections	MIOU 98.1%	<ul style="list-style-type: none"> Lacks multiple diseases on leaves and segmentation of multiple leaves Lacks class (disease)-wise analysis of the results
Feng et al. (2023)	Own in-field images	Alternaria, grey, brown, and rust spot on leaves	Introducing new pooling method to segmentation models with the best result acquired by the ResNet50+DeepLabv3+ model	PA 99.07%, MIOU 82.1%	<ul style="list-style-type: none"> Lacks multiple diseases on leaves and detection of multiple leaves Lacks disease differentiation
Agarla et al. (2023)	Kleynen et al. (2005)	Multispectral images of apple defects variants: russets, scar tissue, frost damage, scald, hail damage, limb rubs, and visible flesh damage	<ul style="list-style-type: none"> Data synthesis method for image augmentation A U-shaped CNN architecture with noise reduction blocks 	F1-score 78% for RGB+NIR images, 79% for RGB images	Lacks ablation study to show the effect of model modifications
Wang et al. (2024)	Own (indoor and outdoor images)	Alternaria blotch, brown spot, grey spot, and rust	U-Net with integrated Residual CBAM attention module in decoder stages and ASPP in the bridge	IOU for leaf 98.35%, for disease 87.15%	<ul style="list-style-type: none"> Lacks disease differentiation Single leaf with single disease were considered
Zhang et al. (2023b)	Plantvillage + own images	Alternaria blotch, brown spots on leaves, as well as multiple diseases (including brown spot and mosaic)	U-Net model with combined loss functions (Dice and Focal loss) and integrated attention module (CBAM)	MIOU 91.07%, MPA 95.58%, F1-score 95.16%	<ul style="list-style-type: none"> Diseases spots on multiple leaves, were not considered False detection part between the leaf and the background in shadow, leaf occlusion and folding
Liu et al. (2022a)	Own in-field images	Alternaria leaf blotch	A two-stage segmentation: MobileNetV2 + PSPNet for leaf segmentation and VGG + U-Net for disease segmentation	MIOU 98.42% for leaf, 92.05% for disease	<ul style="list-style-type: none"> Only a single leaf was considered Mis-segmentation of small lesions and leaf margins
Ding et al. (2025)	Own in-field images	Brown spots, rust, grey spots, and healthy leaves	Integrating a fusion module combining different attention modules (Coordinate Attention, ECA Attention, CBAM, and Triplet Attention) and different dilation rate in ASSP module into DeepLabV3+	MIOU 98.00%, MPA 98.95%, precision 98.45%	Single leaf segmentation

Table 4

Comparison of techniques in plant tissue and disease segmentation: Traditional to CNN and Transformer-based methods.

Category	Method	Advantages	Disadvantages
Traditional	Threshold-based	<ul style="list-style-type: none"> Easy to implement and fast with low computational costs Works with limited images 	<ul style="list-style-type: none"> Sensitive to colour, saturation, and light variation, Affected by background colour Struggles on boundary separation in complex backgrounds Require pre- and post-processing to enhance the results Requires appropriate number of clusters and threshold values to cover colour variations in different images Low generalisation in new images
	Colour-based		
	Clustering-based		
CNN-based	U-Net	<ul style="list-style-type: none"> High performance in more localised area such as diseases Flexible structure to change the model complexity Works well with moderate amount of data 	<ul style="list-style-type: none"> Requires labelled data Limited global context and loss of spatial information Computationally expensive due to several conventional and deconvolutional operations
	SegNet		
	DeepLabV3+		
	PSPNet	<ul style="list-style-type: none"> Lighter model compared to U-Net and DeepLabV3+ Good performance in large objects such as apple or leaf segmentation 	<ul style="list-style-type: none"> Poor boundary segmentation Lack of localised information by pooling layers
Transformer-based	SegFormer	<ul style="list-style-type: none"> Captures global contextual information 	<ul style="list-style-type: none"> Can overlook the small regions like areas at the leaf tips Requires larger datasets Slower inference

Table 5
Publicly available apple disease datasets.

Dataset	Ref	Features	Annotation	Limitation
Plantvillage	Hughes et al. (2015)	<ul style="list-style-type: none"> • RGB images with simple background • Scab, rot, cedar apple rust on leaves 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Low resolution for disease segmentation • Lacks in-field images
Plant pathology2021 (An improved version of Plant Pathology 2020)	Thapa et al. (2021)	<ul style="list-style-type: none"> • RGB and in-field images • Frog eye, powdery mildew, rust, scab • Single and multiple diseases on leaves 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • No specific label for images with multiple diseases (labelled as <i>complex</i>)
PlantDoc	Singh et al. (2020)	<ul style="list-style-type: none"> • RGB images with uniform and complex background • Scab and rust on leaves 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Low number of images • Wrong annotation in some images
Apple leaf disease	Son (2021)	<ul style="list-style-type: none"> • RGB images with complex background • Marssonina blotch and Alternaria on leaves 	<ul style="list-style-type: none"> • Bounding box for object detection 	
ATLSDS	Feng and Chao (2022)	<ul style="list-style-type: none"> • RGB images with simple and complex background • Alternaria spot, grey spot, brown spot and rust on leaves 	<ul style="list-style-type: none"> • Segmentation 	<ul style="list-style-type: none"> • Inaccurate annotation with overlooked small lesions
Apple dataset	Khan et al. (2022a)	<ul style="list-style-type: none"> • RGB images with simple and complex background • Scab, Alternaria spot, powdery mildew, Marssonina leaf blotch, necrosis, insect damage and multiple diseases on leaves 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Only a limited subset of dataset is publicly available
Indigenous dataset for apple leaf disease detection and classification	Yatoo and Sharma (2024)	<ul style="list-style-type: none"> • RGB in-field images • Mosaic and Alternaria spot on leaves • Augmented images 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Low resolution
Kashmiri apple plant disease dataset	Sharma et al. (2023)	<ul style="list-style-type: none"> • RGB images • Apple rot, Alternaria spot and scab on leaves 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Without in-field images
AppleScabLDs & AppleScabFDs	Kodors et al. (2021)	<ul style="list-style-type: none"> • RGB images • Scab lesion on leaves and fruits 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Images are from only one disease type
AL9EE	Zhu (2023)	<ul style="list-style-type: none"> • RGB images with simple and complex background • Alternaria spot, brown spot, frog eye, grey spot, mosaic, powdery mildew, rust, and scab on leaves • Generated through data augmentation using CycleGAN 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Some images are in low resolution for disease segmentation • Some images are incorrectly annotated
Apple dataset	Zhang et al. (2023a)	<ul style="list-style-type: none"> • RGB in-field images • Ring rot and anthracnose on fruits 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Only a limited subset of dataset is publicly available
NZDLPlantDisease	Saleem et al. (2022)	<ul style="list-style-type: none"> • RGB in-field images • Black rot on fruit, black spot (scab) on leaves, European canker, Glomerella leaf spot 	<ul style="list-style-type: none"> • Bounding box for object detection 	<ul style="list-style-type: none"> • Low resolution for disease segmentation, especially for small lesions
ERWIAM	Maß et al. (2024)	<ul style="list-style-type: none"> • RGB in-field images • Fire blight disease 	<ul style="list-style-type: none"> • Bounding box for object detection 	—
Near infrared hyperspectral dataset of healthy and infected apple tree leaves images for the early detection of apple scab disease	Nouri et al. (2018)	<ul style="list-style-type: none"> • Images of infected and healthy apple leaves using NIR hyperspectral camera in laboratory 	<ul style="list-style-type: none"> • Classification 	<ul style="list-style-type: none"> • Image collected in an controlled environment may not generalise to field conditions

(continued on next page)

4.1.3. Gaps

- **Diseases on different parts:** While most recent segmentation studies focused on apple leaf disease detection, there is a need to detect disease both on fruits and stems, such as that for European canker, apple scab and rot. [Saleem et al. \(2022\)](#) address this gap through object detection techniques; however, there is still a need to provide a pixel-wise prediction through semantic

segmentation. Accordingly, collecting a combined dataset in both controlled and uncontrolled environments, on diseases that occur on different plant parts, and under various lighting conditions may contribute to robust prediction.

- **Early to severe and single to multiple disease segmentation:** The diagnosis of diseases from different stages has not received much attention, even though including images from different

Table 5 (continued).

Dataset	Ref	Features	Annotation	Limitation
Visible-near infrared hyperspectral dataset of healthy and infected apple tree leaves images for the monitoring of apple fire blight	Gaci et al. (2023)	• Hyperspectral images of apple tree leaves infected with fire blight	• Classification	• Images are from only one disease type

stages could improve generalisations and reliability. As shown in Fig. 16, young lesions have different colours compared to mature lesions in some apple diseases like frog eye or scab. In particular, this matters in the process of severity assessment tasks. Furthermore, the segmentation of multiple diseases needs to be considered more in semantic segmentation studies. Lastly, regarding the annotation phase, more research needs to be conducted to show how the annotation method could affect the results of segmentation. We have started to address above-mentioned gaps by considering both apple fruits and leaves, multiple diseases, as well as assessing the effect of annotation in our recent publication (Keshavarzi et al., 2024).

• Robust augmentation

- Data augmentation remains a valuable tool in addressing limited data availability and enhancing model generalisation. Various augmentation techniques need to be refined to generate diverse and realistic training data while preserving the necessary details for agricultural analysis. Weather variability may cause illumination changes, dust, raindrops, or shadows on plants. Therefore, in-field image augmentation needs much more work to imitate these parameters rather than simply adding blur and noise filters. Additionally, a complete assessment of the effects of different augmentation methods on the segmentation process is required.

- Data augmentation through GANs has demonstrated significant usefulness in the case of different applications of semantic segmentation in agriculture, including leaf counting, fruit detection, etc. However, the output generated by GAN techniques may not capture fine details in texture and colour, which are essential for accurate disease segmentation. The effect of data augmentation using GANs on each type of apple disease and its impact on the results could be investigated.

- The use of synthesised images to recreate future disease stages could be a solution for cross-stage robustness in apple disease detection. Attempts have been made to enhance model generalisation by synthesising future growth in some crops (Drees et al., 2021), but it remains unresolved in apples.

- **Domain-specific transfer learning:** Many studies commonly utilise backbone models pre-trained on general domain datasets, such as ImageNet. However, there remains a need to investigate the impact of employing backbone models trained on related domain datasets, such as PlantCLEF2023 (Goëau et al., 2023), specifically in the context of apple disease segmentation. Lee et al. (2020) observed that while fine-tuning a (CNN) model on domain-related datasets like PlantCLEF could reduce overfitting, this does not necessarily contribute to enhancing generalisation in plant diseases classification. This underscores the importance of exploring the effectiveness of domain-specific pre-training for improved performance in apple disease segmentation tasks.

4.2. Model development

4.2.1. Challenges: Boundaries, small objects, imbalanced data

Based on the reviewed literature, complex backgrounds is one of the main challenges in agricultural images for semantic segmentation. The background often contains cluttered elements such as soil, weeds, and overlapping leaves which, can cause false predictions, especially

at boundaries. Furthermore, some studies reported poor detection of occluded leaf or fruits, as well as small apples or diseases.

In plant disease segmentation, datasets are often imbalanced, and data distribution is not necessarily influenced by the number of images in each category. This is because a single image might include most labels, while others may only exhibit a few. Consider an image with the *rust* label, which consists of six leaves with two lesions each. Alternatively, another image might have the label *scab*, featuring two leaves but with twenty lesions. Some labels in the dataset may have a minority of pixels compared to *background* which includes the majority of pixels in an image and thus the entire dataset. Various studies have adopted different strategies to address these issues, some of which are highlighted below.

4.2.2. Solutions

4.2.2.1. Model modification. Li et al. (2021) proposed a modified U-Net model to tackle the weak boundaries in segmentation of green apples from a complex background. To effectively extract edge information, residual blocks and gated convolutions were used to form an additional branch to extract information from the downsampled layers. Results showed a 5.75% increase in the F1-score for apple fruit segmentation.

Sheng et al. (2023) added the edge information to a region-based module to detect the precise region of apples with a complex background. This method performed well in separating green apples from a green background. Global localisation information was derived through a multi-scale localisation module to extract features with different kernel sizes. By integrating the feature information of high and low layers, the boundary-aware module sharpened the edges of potential targets. Although apples are segmented successfully using this method, segmentation of small apples was more complicated.

Pooling layers as one of the methods to select prominent values, usually uses the max or average operation with $n \times n$ kernel sizes. Strip pooling was introduced to the pooling operation along the horizontal and vertical axis with a kernel size of $1 \times n$ or $n \times 1$ to capture long-range dependencies (Hou et al., 2020). Feng et al. (2023) expanded the pooling method to employ strip and twill pooling operations to propose the double cross pooling method. The double cross pooling module does the horizontal pooling, vertical pooling, primary diagonal twill pooling and secondary diagonal twill pooling, respectively. The fused results add weights to the feature map through multiplication. Integrating this pooling module into DeepLabV3+ improved the disease segmentation on apple leaves by 7.69%.

Zhu et al. (2023b) argued that the MLP-based decoder in SegFormer is not efficient in segmenting small objects against complex backgrounds while Transformers look at a larger scale. They proposed modifications to SegFormer to segment the small apples, whereby two modules are introduced in the decoder part. One module designed to enhance the segmentation quality by updating convolutional kernels during the upsampling process, and the second module, Complex Points, refines masks with less blurred edges by optimising the segmentation of hard-to-classify points. This method outperformed other models such as DeepLabV3+ and FCN in segmenting small fruits on apple trees. On the other side, the Segformer model showed lower MIOU in indoor images having apple rust disease compared to the CNN-based models, DeepLabV3+ and U-Net, by 0.59% and 0.71% (Lu et al., 2024).

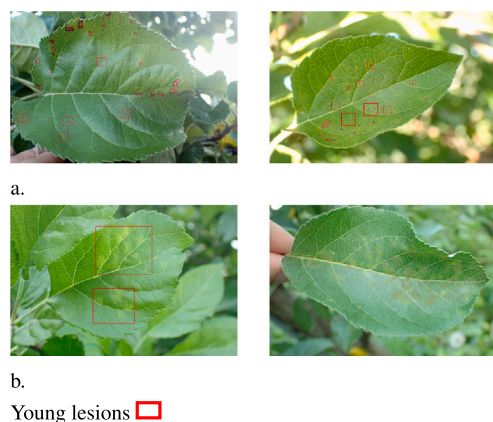


Fig. 16. Young and mature lesion differences associated with some apple leaf diseases. a. Leaf with frog eye disease. The young purple lesions have a different appearance to mature ones. b. Pale young scab lesions compared to a more mature olive-coloured spot with a velvety texture.

Source: Images from Thapa et al. (2021).

Lu et al. (2024) developed EAIS-Former with a dual overlap patch embedding method. This model benefited from combining the advantages of transformers and CNNs. This method uses two different patch sizes of 3×3 and 5×5 to capture the lesions on mango and pomegranate leaves, and also apple rust disease at different scales. The self-attention in transformers is then replaced with a residual block consisting of two depthwise layers with a 5×5 kernel and a large kernel (near the feature map size) to learn both location and global information. Their feature extraction and two-stage upsampling method preserved the details, providing a more precise segmentation of small apple rust, grey, and Alternaria spots with an MIOU of 99.04%, 98.58%, and 98.59% respectively.

Lu et al. (2023) designed three modules in an encoder-decoder shape segmentation network, inspiring transformers and CNN. Mix transformer, position indexer and local optimisation modules were integrated to establish global dependencies, preserve the position of information in down-sampling, and enhance the local and detailed feature extraction, respectively. While, according to Mzoughi and Yahiaoui (2023) one major failure in segmentation relates to lesions which appear in leaf ribs and margins, their method led to more precise segmentation of apple grey spots in leaf regions and small Alternaria spots in occluded leaves.

4.2.2.2. Attention mechanism. In plant disease segmentation, features extracted from critical areas like leaves and lesions are more informative than features from the background. When multiple leaves and lesions appear in different locations of an image, global information may lead to a better capture of the diseases globally. With their self-attention mechanism, transformers have the advantage of capturing global dependencies between different parts of an image. However, they have more complexity and are less sensitive to positional information. Some studies tend to direct models to emphasise the region or area of interest in the image by giving more weight to the most relevant information while ignoring less important details. It could effectively give more attention to the leaf and disease areas or focus on small lesions. While it is a broad topic in DL models, we will outline some implemented methods in apple and other plant disease segmentation.

Attention U-Net (Oktay et al., 2018) uses an attention gate at the top of each skip connection in the U-Net model (illustrated in Fig. 17). The attention gate calculates the attention coefficient values to actively suppress irrelevant regions and reduce the semantic gap between the encoder and decoder. A $1 \times 1 \times 1$ convolution layer performs the linear transformation and computes a weighted combination of the channel values at a specific location. The output of the attention gate is the

element-wise multiplication of input feature maps and attention coefficients. More relevant features will be more involved in upsampling and final classification. The attention gate shows promising results in segmentation of small-sized disease lesions on for example, rice (Rai and Pahuja, 2023), maize, and cucumber (Zhang and Zhang, 2023), by reducing background intervention. Xu et al. (2020) used an attention module with two levels to alleviate background effects and highlight spots. Inter-layer and intra-layer attention weights were derived which showed good results in separating rice plants from a very similar background to the rice plants.

Squeeze-and-Excitation (SE) Networks (Hu et al., 2020) introduced an attention module which has been used in segmentation models for detection of diseases in tomatoes (Deng et al., 2023), rice (Chen et al., 2021), and apples (Zhang et al., 2023b), yielding improved results. Fig. 18 illustrates the SE attention module that reweighs each channel through global pooling, fully connected layers, and activation functions.

Yang et al. (2024) modified the U-Net model by integrating the SegFormer in the encoder, efficient channel attention (ECA) (Wang et al., 2020) at the bridge, and a multi-scale fusion attention block (MFAB) in the decoder part. ECA is a lightweight and enhanced version of the SE attention method, replacing the FC layer with a 1×1 convolution layer. This allows it to capture inter-channel dependencies among K channels through a local cross-attention module, with the K value adaptable as needed. MFAB incorporates two SE attention modules: one for high-level features and another for low-level features. The final model had fewer parameters and resulted in a slight improvement in segmentation of grey, brown, and rust disease spots in apple leaves (0.2%).

The convolutional block attention module (CBAM) (Woo et al., 2018) comprises both channel and spatial attention modules (Fig. 19). The channel module, like in the SE attention module, emphasises the importance of each specific channel, but with both max-pooling and average-pooling operations. The spatial attention module performs the pooling (max and average) and a convolution operation which results in a 1-channel feature map ($1 \times h \times w$) to assign weights to different spatial locations. Focusing on small lesion segmentation, Zhang et al. (2023b) integrated different attention modules into the U-Net model to segment the Alternaria spots and brown spot diseases on apple leaves. Results showed the SE attention module improved results compared to the U-Net model, while the CBAM module improved the segmentation of small Alternaria spots. Introducing the CBAM module into DeepLabV3+ increased the segmentation results in apple leaves spots by 7.53% (Li et al., 2023b). In essence, channel attention involves looking for *what* are the significant across all channels of a feature map refining the relevant features accordingly. On the other hand, spatial attention determines the significant locations within the image, emphasising *where* these important elements are situated. Studies showed that combining these two types of attention mechanisms could improve the recognition of regions or areas of interest in plant disease segmentation (Xu et al., 2020; Cai et al., 2022). In a field-based case study by Xue et al. (2023) in which precise spraying by orchard spraying robots was enabled, the DeepLabV3+ was improved by integrating CBAM into its ASSP module for fruit tree canopy segmentation. MobileNetV3-Small was selected as the backbone network to make the model lighter and six times faster. The model was integrated into the NVIDIA Jetson Xavier NX device for efficient on-site inference with an MIOU of 95.62%.

4.2.2.3. Loss function. While data augmentation and transfer learning techniques could alleviate the image scarcity problem, it is important to address imbalanced datasets in segmentation tasks, as simply increasing the number of images may not fully solve this issue. One solution involves utilising appropriate loss functions during training to focus on the hard samples (low-frequency labels). In DL models, the loss or cost function quantitatively measures the difference between prediction and actual ground truth labels. The goal is to minimise the cost value by comparing the predicted and true labels of validation data during

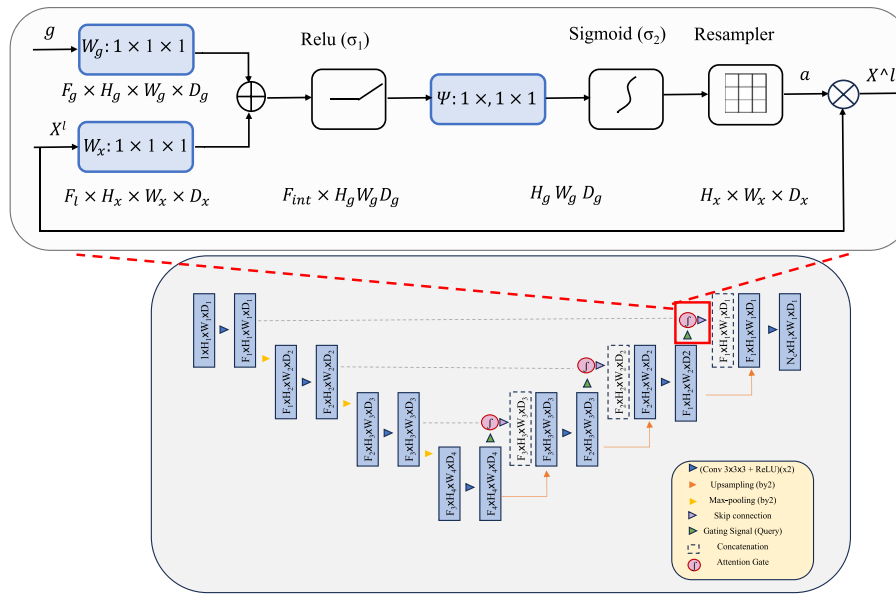


Fig. 17. Architecture of attention U-Net model. Source: Image from Oktay et al. (2018).

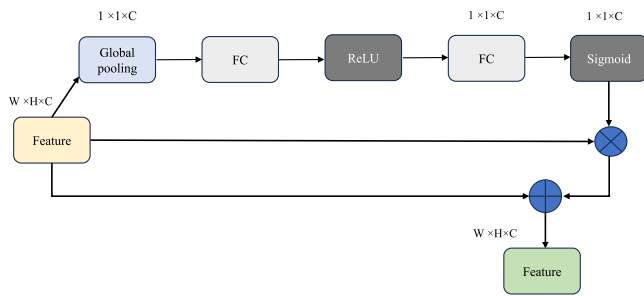


Fig. 18. Block diagram of Squeeze-and-Excitation module. The feature map goes through two fully connected layers with a ReLU activation function in between. The sigmoid activation function provides a simple gating mechanism and a vector of $c \times 1 \times 1$ weight for channels. Source: Image adapted from Chen et al. (2021).

training. The categorical cross-entropy (Ma et al., 2004) is a standard loss function which is widely used in many works (Roy et al., 2021; Lian et al., 2024; Storey et al., 2022; Bedi et al., 2024; Sheng et al., 2023; Khan et al., 2022b; Xu et al., 2020). This loss function is defined as:

$$L(y, \hat{y}) = - \sum_i y_i \log(\hat{y}_i) \quad (2)$$

where y is the one-hot ground truth and \hat{y} is the prediction distribution. Focal cross entropy (Lin et al., 2017) is one of the loss functions introduced for highly imbalanced data sets which results in predicting hard samples (Zhang et al., 2023a; Liu et al., 2024a) more precisely. Categorical focal loss is a modified version of the standard categorical cross-entropy loss function. This function introduces a focusing parameter γ and a modulating factor $(1 - \hat{y}_i)^\gamma$ for each class. The modulating factor downweights the well-classified samples with γ as a rate.

$$L(y, \hat{y}) = - \sum_i y_i (1 - \hat{y}_i)^\gamma \log(\hat{y}_i) \quad (3)$$

To address the problem of small apple lesion segmentation and disease similarity at an early stage, a combination of focal loss and dice loss was integrated into U-Net model (Zhang et al., 2023b). This increased the MIOU by 0.62% compared to using cross-entropy loss function. The

dice function is defined as below:

$$\text{Dice Loss} = 1 - \frac{2 \times \sum (y_{\text{true}} \times y_{\text{pred}}) + \epsilon}{\sum y_{\text{true}} + \sum y_{\text{pred}} + \epsilon} \quad (4)$$

\sum denotes the sum over all pixels in the masks. ϵ is a small constant (usually added to the denominator) to prevent division by zero, ensuring numerical stability. $y_{\text{true}} \times y_{\text{pred}}$ is the element-wise multiplication (intersection) of the ground truth and predicted masks.

Agarla et al. (2023) employed the Focal Tversky Loss (FTL) (Abraham and Khan, 2019) to mitigate the impact of high-frequency pixels from healthy regions compared to the smaller number of pixels in defect areas of postharvest apples. This function is defined as follows:

$$\text{TI}_c = \frac{\sum_{i=1}^N p_{ic} g_{ic} + \epsilon}{\sum_{i=1}^N p_{ic} g_{ic} + \alpha \sum_{i=1}^N (p_{i\hat{c}} g_{ic}) + (1 - \alpha) \sum_{i=1}^N (p_{ic} g_{i\hat{c}}) + \epsilon} \quad (5)$$

$$\text{FTL}_c = \sum (1 - \text{TI}_c)^{\frac{1}{\gamma}} \quad (6)$$

where, in their problem with labels (lesion and healthy), p_{ic} represents the predicted probabilities for each pixel i in lesion class c , $p_{i\hat{c}}$ is the healthy probability of the pixel i . g_{ic} and $g_{i\hat{c}}$ are the ground truth label c and healthy label \hat{c} for each pixel i .

Zhu et al. (2023a) proposed an adaptive loss function based on the cross-entropy loss function and a modulation factor for a two-stage apple leaf disease segmentation from a complex background.

$$\text{AL}(y, p) = \begin{cases} -(\cos(p + \frac{\pi}{2}) + 1)^\beta \log(p) & \text{if } y = 1 \\ -(\cos(1 - p + \frac{\pi}{2}) + 1)^\beta \log(1 - p) & \text{if } y = 0 \end{cases} \quad (7)$$

$\beta \in [0, 5]$ is a hyperparameter showing the compression degree that needs to be set experimentally and $[\cos(1 - p + \frac{\pi}{2}) + 1]$ is determined by p (the probability of each pixel belongs to class c). In easily classified pixels, the higher the probability, the lower the modulation factor, and the greater the compression. Their loss function improved results in segmenting overlapped leaf areas and small lesions missed with the standard cross-entropy loss function (0.11% and 1.27% for leaves and disease spots respectively).

Another technique that gives more attention to the least common labels involves giving weight to each label and making the model learn all data with the same weights (balanced weights) or giving desired weights to labels depending on the application. Wang et al. (2024) proposed a weight compression loss function to decrease the effect of background labels and give more attention to the hard-to-categorise

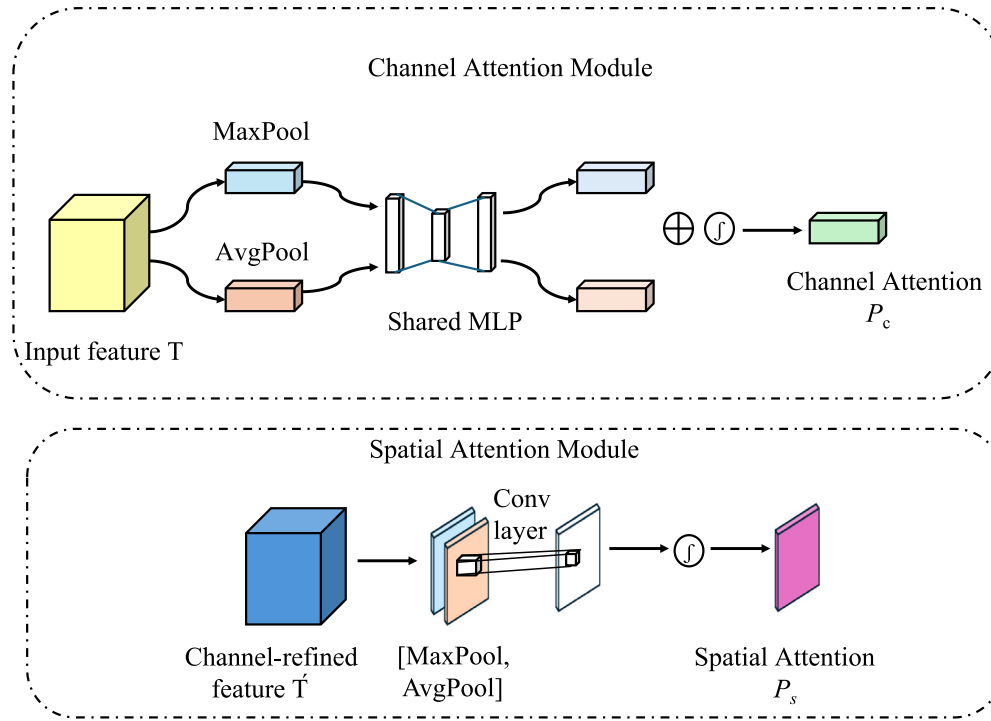


Fig. 19. Block diagram of CBAM. Both max-pooling and average-pooling operations are employed on the feature maps in both the channel and spatial attention sub-modules. Source: Image from Woo et al. (2018).

lesion labels. The cross-entropy loss function was modified by adding a modulation factor ($[1-\arctan(p)]$) which its best value acquired through different test.

$$WCLce(y, p) = \begin{cases} -(1 - \arctan(p))^\gamma \log(p) & \text{if } y = 1 \\ -(\arctan(p))^\gamma \log(1 - p) & \text{if } y = 0 \end{cases} \quad (8)$$

4.3. Evaluation of the segmentation results

4.3.1. Measurement metrics

In a plant disease detection problem with two classes (for example healthy and diseased), positive detection is detection of disease and negative detection refers to healthy tissue. From the confusion matrix, each image (or pixels for segmentation) falls into one of four categories: **True positive (TP)**: The number of diseased pixels correctly predicted as diseased.

False positive (FP): The number of healthy pixels incorrectly predicted as diseased.

True negative (TN): The number of correctly predicted healthy pixels.

False negative (FN): The number of diseased pixels predicted as healthy.

However, for multi-class problems (including semantic segmentation), the definitions are applied class-wise (see the confusion matrix in Table 6).

Pixel accuracy (PA),

$$PA = \frac{TP + TN}{TP + TN + FN + FP} \quad (9)$$

is one of the most commonly employed metrics for evaluating models in DL classification tasks due to its simplicity. It represents the rates of total correct classified pixels. However, its simplicity might not render it the most effective metric for assessing disease lesion segmentation, particularly in imbalanced datasets. If many pixels are correctly classified in the background, the PA value for a disease lesion or leaf class will increase. PA combines both FP and FN , making it less informative

for analysing segmentation errors. Precision, recall, F_1 -score and MIOU, are other metrics for evaluating the segmentation results. Precision,

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

measures the validity and accuracy of positive predictions, showing to what extent results are reliable. Reducing FP results in higher precision values. Recall,

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

emphasises reducing the FN rate and measures the model's ability to label the actual positive ground truth correctly. It is useful for understanding how portions of diseases or healthy leaves and fruit have been correctly identified. MIOU,

$$mIoU = \frac{1}{N} \sum_{i=1}^N \frac{|P_i \cap G_i|}{|P_i \cup G_i|} \quad (12)$$

measures the average overlap area between predicted and ground truth regions among all pixels where P_i and G_i are the predicted pixels and ground truth pixels, respectively. This measurement has been used widely in assessing the apple disease segmentation results (Yu et al., 2019; Hasan et al., 2022a; Liu et al., 2024a; Mahmud et al., 2023; Li et al., 2023a; Lu et al., 2024; Zhu et al., 2023a). According to the problem type (diseased and healthy prediction), precision and recall are important in assessing the whole model's performance. An increase in precision and recall could result in more reliable results, reducing the cost of treatment, and estimating loss more precisely. Thus, there should be a trade-off between these two metrics. F_1 -score,

$$F_1 \text{ score} = 2 \times \frac{precision \times recall}{precision + recall} \quad (13)$$

as a harmonic mean of precision and recall is commonly used in apple disease segmentation problems dealing with unbalanced datasets (Zhu et al., 2023b; Bedi et al., 2024; Yang et al., 2024). As the defect area is usually a smaller portion of the healthy part of the apple, Agarla et al. (2023) proposed another metric in apple fruit defect segmentation to

Table 6
Confusion matrix in multi-class classification problems.

		Predicted		
		$class_0..class_{k-1}$	$class_k$	$class_{k+1}..class_n$
Actual	$class_0..class_{k-1}$	TN	FP	TN
	$class_k$	FN	TP	FN
	$class_{k+1}..class_n$	TN	FP	TN

measure the Class-Specific Recognition Error (CSRE).

$$CSRE = \frac{FN}{TP+FN} + \frac{FP}{TN+FP} \quad (14)$$

4.3.2. Ablation studies and visualisation

Ablation studies analyse the impact of specific components of a model on the model's performance. This helps to understand the contribution of different elements and assess their significance in achieving the desired segmentation result. In some studies, ablation has been used to show the effectiveness of different attention mechanisms (Wang et al., 2024; Zhang et al., 2023b; Feng et al., 2023; Dai et al., 2024). In another study, Liu et al. (2024a) performed an ablation experiment to show how deformable convolution affects the segmentation of frog eye disease in complex backgrounds. Lu et al. (2024) performed the ablation experiments to show the adaptability of their proposed modules. They also performed a visual comparison of the results with different modules. Visualisation can also illustrate the effectiveness of different receptive fields in convolution layers (Zhu et al., 2023b), display the output of the feature maps after deploying different modules (Feng et al., 2023), and demonstrate feature representation (Mahmud et al., 2023; Yang et al., 2023). Additionally, they can highlight *TP*, *FP*, *TN*, *FN* in prediction results (Liu et al., 2022a; Zhang et al., 2023b; Liu et al., 2024a; Li et al., 2023a; Zhu et al., 2023a; Dai et al., 2024; Jiang et al., 2024).

4.3.3. Gaps

- **Ignored healthy class:** While many studies focus solely on metrics related to disease classes (Mzoughi and Yahiaoui, 2023; Mahmud et al., 2023; Li et al., 2023a; Lu et al., 2024, 2023; Bedi et al., 2024; Feng et al., 2023; Wang et al., 2024; Zhang et al., 2023b), there is a need to extend this evaluation to include the healthy class. Fewer works (Lian et al., 2024; Zhu et al., 2023a) have incorporated metric evaluations specifically for the healthy class. The evaluation of both diseased and healthy classes is essential for comprehensive model assessment in tasks such as disease segmentation. The model must effectively recognise healthy and diseased regions because misclassifications can have serious consequences. Identifying healthy parts as diseased could result in unnecessary treatments or interventions while overlooking diseased regions could result in untreated conditions, both carrying significant costs and implications.
- **Lack of class-wise analysis:** Class-wise results could provide a more in-depth understanding of the model performance. While reporting a total segmentation accuracy or MIOU (Feng et al., 2023) is a good indication of segmentation model efficiency in apple disease detection, they do not reflect the model's ability to correctly distinguish between different classes.

It has been shown that visualisation techniques can improve the interpretation in ablation studies (Wang et al., 2024; Zhang et al., 2023b; Feng et al., 2023; Yang et al., 2024; Zhu et al., 2023a,a). Visualisation can help to make the represented results derived from ablation experiments clearer. Furthermore, utilising only the measuring metrics in ablation studies (Sheng et al., 2023; Bedi et al., 2024; Son, 2021), may not provide detailed insights from the results. On the other hand,

some minor changes might not be visually noticeable and easy to follow, making it difficult to make comparisons based on images. This is where measuring metrics come into play by reflecting subtle differences quantitatively. Hence, by combining all these approaches, researchers can determine how different model components affect performance while visualising and quantifying these effects.

5. Multispectral and hyperspectral imaging

Some fruit defects are difficult to recognise visually and may not be captured effectively through RGB images. In such cases, advanced imaging methods, such as multispectral or hyperspectral imaging, can be more helpful. These techniques, with their sensitivity to subtle colour changes and broad spectral range, can identify decay at early stages in apple tissue. These imaging techniques have been used for early bruise detection. In a study, Tian et al. (2024) employed near-infrared (NIR) imaging and adaptive threshold segmentation for detection of apple bruises. Full-band images were captured from the fruit face, stem end, and calyx end of both healthy apples and those with single or multiple bruises, 30 min after bruise occurrence. The proposed method achieved an F1-score of 94.70%, outperforming other segmentation techniques such as Otsu and fixed thresholding. In another study, a watershed segmentation algorithm was deployed on the most effective spectral regions identified through principal component analysis (PCA) (Luo et al., 2019). The output of the binary segmentation image was used to classify samples into healthy and bruised apples. The same strategy was used in Li et al. (2019) to extract PC scores and segment the decay region on postharvest apples through a watershed algorithm and postprocessing morphological operations. The images were taken using three spectral regions including Vis-NIR (400–1000 nm), Vis (400–780 nm) and NIR (781–1000 nm). The classification accuracy for 220 decayed and 220 healthy fruit were 99% and 100% respectively. Tian et al. (2023) also applied Vis-NIR hyperspectral reflectance imaging and developed a bruise detection algorithm using optimal band ratio images and pseudo-colour transformation, followed by an improved watershed segmentation algorithm. They achieved recognition rates of 93.3% for healthy and 92.2% for bruised apples. The segmentation results then can be used in classification tasks for grading and quality evaluation of postharvest fruits. However, these imaging technologies need complicated equipment and preprocessing techniques to extract the most effective regions, and their practicality still needs to be tested on an online system for realtime detection of apple bruise and decay. Furthermore, a deeper exploration of the effectiveness of these imaging modalities compared to RGB images is needed, as Li et al. (2019) suggested that segmentation results in postharvest apple defect detection from multispectral and RGB images are comparable.

Apart from segmentation tasks, there have been some studies focusing on classification and identification of apple diseases using hyperspectral and multispectral imaging. For instance, Bleasdale and Whyatt (2024) used RGB and NIR images to classify apple leaves with invisible (early stage) and mature scab infections. EfficientNetV2L model resulted in 97.57% accuracy over the combined dataset and 78.91% over the multispectral dataset. In another study, Delalieux et al. (2009) utilised indices derived from hyperspectral and fluorescence images to detect early scab infections before visible symptoms appeared under a controlled environment with favourable water and soil moisture conditions. Vegetation indices derived from multispectral images have also been effective in identifying healthy and unhealthy canopy sections infected with diseases like powdery mildew (Chandel et al., 2020) and fire blight (Mahmud et al., 2023; Xiao et al., 2022). These indices were subsequently used to classify entire trees or canopies into healthy or unhealthy categories. Furthermore, Skoneczny et al. (2020) identified the most effective hyperspectral bands for differentiating apple leaves as healthy, infected with fire blight, or dry. These indices can be employed for the real-time automatic detection of the

diseases. Hyperspectral imaging was also applied to study apple mosaic disease, comparing spectral differences between healthy and infected leaves (Jiang et al., 2023). Although these studies primarily focused on disease identification, the derived indices and measurements, such as spectral reflectance and vegetation indices, can also be adapted for segmentation tasks. Techniques like thresholding can leverage these indices to segment diseased and healthy regions.

6. Apple disease severity assessment

Identifying the infected area (disease incidence) can be a preliminary step in assessing disease severity by determining what portion of the plant, such as specific leaves or the entire orchard, is infected. It is useful for developing management strategies, including yield loss prediction, estimation of when and how much pesticide to apply or spray, and evaluating their impacts on diseases (Bock et al., 2022). In visual disease severity estimation three types of scales are commonly used: Nominal or descriptive scales, ordinal scales (quantitatively and qualitatively), and ratio scales.

- Nominal scales rely on concise descriptions like “Healthy”, “Mild”, “Moderate”, and “Severe” to show various levels of disease severity.
- In quantitative ordinal scales, numeric intervals between 0% and 100% are described by a set number of classes. Pal and Kumar (2023) benefited from the percentage of the diseased area to label their samples with nominal scales (0% = Normal, <20% = Mild, ≥20% and <30% = Moderate, ≥30% = Severe). Qualitative ordinal scales usually have no specific and known numeric magnitude between each class, which is useful in estimating the severity of diseases that cannot be readily quantified (Bock et al., 2020).
- The percentage scale, is a commonly utilised method for visually estimating the level of disease severity (Bock et al., 2020). Numerous disease intensities can be assessed using the ratio of the diseased area to the whole area (Bock and Chiang, 2019; Hamada et al., 2019; Xue et al., 2019).

Visual assessment of diseases by raters is prone to subjectivity, requiring rater training and testing. Moreover, standard area diagrams (SADs) are usually needed as an aid for the assessments (Bock et al., 2022, 2020; El Jarroudi et al., 2015). Recently, DL models and image analysis-based methods demonstrated superiority in disease severity assessment by providing more accurate and reliable results (Sun et al., 2014). Some researchers classified image datasets into different nominal classes to predict the severity level of test samples (Wang et al., 2017; Liang et al., 2019; Prabhakar et al., 2020; Hayit et al., 2021; Dhiman et al., 2022; Ji et al., 2020). The nominal scales are subjective and can vary according to different raters. Moreover, these classification models may lead to incorrect severity assessment in cases of new samples with varying levels of diseases (Shi et al., 2023). Determining the diseased and healthy area is the key element in plant disease severity assessment. Therefore, segmentation is a strong tool that can assist in more accurate disease and leaf separation to measure the disease intensity.

There are some tools developed for disease severity measurements, although studies showed that only a limited number of them provide both disease detection and severity assessment (Siddiqua et al., 2022). The CompuEye (Bakr, 2005), Assess (Lamari, 2002), and ImageJ (Schneider et al., 2012) tools utilise colour-based and thresholding methods to segment the disease area. In 2013 Phenotic software (Rousseau et al., 2013) was developed to provide a threshold-based segmentation and disease severity assessment. Bean plants were photographed using a chlorophyll fluorescence imaging system that allowed symptoms to be detected before they appeared visually. Pethybridge and Nelson (2015) released the Leaf Doctor, a semi-automatic software to segment the leaf and disease area and assess the disease severity of individual leaves. In this work, the ability to select eight

green colour ranges in leaves, different colour ranges of symptoms, and leaf veins results in less segmentation error. The interactive nature of this software allows the user to define a threshold showing the maximum number of healthy and unhealthy pixels, especially on margins. However, it is limited to having a solid black background as, otherwise, there would be misclassification of pixels. Above-mentioned tools usually cannot overcome colour variation coming from different stages of disease such as frog eye or scab in RGB images. Esgario et al. (2022) developed an Android app for biotic stress identification and quantification in coffee leaves using PSPNet and U-Net models. Diseases were segmented from a simple white background to measure the severity using a ratio scale. From a subjective assessment, PSPNet showed more smooth and rounded edges than U-Net, which may result in under-segmentation. Liu et al. (2022a) leveraged PSPNet and U-Net to segment apple leaf and *Alternaria* spots, respectively. The obtained ratios of disease severity were classified into five grades, namely healthy (0), early (0–0.95%), mild (0.95–1.75%), moderate (1.75–3.00%) and severe (3.00%–100%). Comparing the manually labelled images and the severity classification showed promising results in application of CNN based segmentation for severity assessment. Some studies used the China Ministry of Agriculture’s criteria to grade apple rust and ring rot on leaves (Feng et al., 2023; Li et al., 2023a) (grade 0 for no disease, grade 1 for the lesion area accounts for less than 10% of the whole leaf area, grades 3, 5, 7 and 9 for lesions covering 11%–25%, 26%–40%, 41%–65% and more than 65% of the whole leaf area, respectively).

Although segmentation is a practical tool for acquiring the diseased portion of plant tissue, it is not solely adequate for defining the visual disease severity. Xing et al. (2023) introduced two parameters to incorporate the distribution of apple leaf spots in severity assessment. Results improved disease severity classification of samples from the PlantVillage dataset with black rot, scab and rust diseases. Hence, precise leaf and lesion segmentation in combination with hand-crafted features of colour, shape and texture, could be applied to achieve better results. Filling some other gaps mentioned in the dataset section such as looking at diseases at larger scales can make the severity assessment through semantic segmentation more robust.

7. Costs of semantic segmentation

Different methods, ranging from conventional to advanced semantic segmentation approaches, have varying levels of data and computational costs. These factors are referred to as a digitisation footprint (Huang et al., 2024b). Considering these costs is crucial for selecting the most appropriate method based on the application and context.

7.1. Data collection effort

Conventional segmentation methods require less effort for data collection, as these algorithms can often be parameterised and applied to a single image. However, for analysing diseases at different stages or detecting multiple diseases, it may be necessary to process multiple images to optimise parameters such as thresholds or cluster numbers, particularly for images with higher variabilities. Supervised DL models, on the other hand, demand large datasets, often thousands of annotated images, to achieve convergence. This significantly increases the time and effort required for data collection. For applications involving multiple disease stages or detecting diseases on both fruits and leaves, multiple rounds of sampling may be necessary. For example, in the case of the in-field apple disease dataset, Plant Pathology (Thapa et al., 2021), two years of sampling were required to adequately capture these complexities. Similarly, the image collection process for the CASC IFW dataset (Singh et al., 2011) also spanned two years. For fire blight symptoms at different stages images were collected between 2021 and 2024, on nine dates under different weather conditions (Maß et al.,

2024). To acquire images under real weather conditions at different stages of disease development, sampling can be undertaken at different times of the day or even over different months, as carried out by Kodors et al. (2021), Zhang et al. (2023a) and Sharma et al. (2023). This process generally requires more effort for DL models compared to conventional methods as they need more data either through image collection or data augmentation (Zhu, 2023). Using multimodal sensors, such as combining RGB and multispectral imagery, further adds to data collection costs in terms of time, labour, and equipment. As mentioned before, multispectral imaging requires advanced equipment and specialised training compared to the relatively straightforward use of RGB cameras.

7.2. Data preparation and processing

Supervised CNN and transformer-based models require extensive image annotation and augmentation, which involve significant time, labour, and computational resources. For traditional segmentation techniques, image preprocessing methods such as enhancement and noise removal are typically used to prepare data. While these methods are less costly in terms of time and processing compared to DL models, they are often limited in their adaptability to new domains. While data preparation for DL models is costly, they are more generalisable to the new domain or new environment (e.g. leveraging transfer learning with pretrained models on large datasets to train a new model for a new plant).

7.3. Computational costs

Computational costs in semantic segmentation methods are usually measured through three different parameters:

- Inference time (ms/image), which represents the time required for the prediction of a single image using a specific method.
- Floating point operations (FLOPs) (mostly GFLOPs/image), which indicate the number of arithmetic operations a model needs to process one image.
- Number of model parameters refers to the total trainable weights in the model. This impacts memory requirements and can influence both training and inference efficiency.

Traditional methods generally have lower computational costs than DL-based models. Tian et al. (2024) compared DL models with conventional approaches and showed that compared to the fastest DL model evaluated in the study (186.70 ms), the thresholding method achieved a lower average detection time per NIR image (131.10 ms). The GrabCut method, however, had a higher segmentation time per image compared to U-Net and SegNet model inference on CPU (Ngugi et al., 2020). Therefore, factors such as the number of training samples, hardware configuration, and image resolution can influence inference time and computational cost, leading to discrepancies across different studies. For example, in two studies U-Net model reported to have different number of parameters (Esgario et al., 2022; Lu et al., 2023). Different model backbones and architectures may cause this. In Esgario et al. (2022) U-Net model had more parameters, but it converged faster with lower inference time compared to PSPNet model. Therefore, comparing methods based on computational cost should be conducted using the same dataset and hardware configuration to be fair.

While conventional methods are more resource-efficient and suitable for systems with limited computational power, deep learning models, despite their higher initial training costs, excel in handling complex datasets and generalising across domains, often making them the preferred choice. However, to benefit from their advantages, their high computational costs needs to be alleviated. While in most studies, researchers make model architectures more complex to reach the optimum results, there is a need to study methods that simplify these

models. For example, one of the model variation introduced by Lu et al. (2023) is a lightweight model with 1.90 M parameters, and 4.17 GFLOPs that can be integrated in mobile devices in leaf disease diagnosis. Liang et al. (2022) leveraged BiSeNetV2 (Yu et al., 2021) to decrease the computational cost and make the model simpler for real-time grading of defect apples with lower parameters (with 2.22 MB parameters). However, results still needed to be improved due to false positives in stem/calyx regions through an object detection algorithm. To reduce computational costs in terms of training time for models to converge, low-cost image processing techniques and more efficient data preparation techniques may be incorporated.

8. Conclusion and future direction

Semantic segmentation is a powerful tool that plays a crucial role in identifying the location, type, and area of plant diseases. This review delves into the evolution from conventional methods to state-of-the-art segmentation techniques employed in plant disease detection, focusing particularly on apples. It also addresses the various challenges encountered and presents implemented solutions at different stages of the apple disease detection and severity assessment process.

While results of studies showed the significance of CNN and transformers models in providing more accurate segmentation, there are still open issues in enhancing results by a better understanding of the whole image. In most recent studies, a single label is allocated to all possible diseases on the leaves (Liang and Gao, 2024; Ding et al., 2025; Lu et al., 2024). There may be a solution in leading models to disease-wise identification. However, using more labels may cause the lower performance. Combining a classification model with segmented lesions could then classify the lesions into disease types. Huang et al. (2024a) used the vice versa method to classify images into disease types, then segment lesions. Therefore, we argue that expanding the idea of *detection* to *identification* is a major step towards more robust models. On the other side, despite the progress made in apple disease segmentation, distinguishing between certain types of diseases such as *Alternaria* leaf spot and frog eye remains a challenge due to the high visual similarity of their characteristics. In this case, utilising multimodal data such as RGB and microscopic images alongside text descriptions could be explored. Additionally, considering the distribution of diseases within a canopy is crucial, as *Alternaria* disease typically exhibits a more uniform distribution than frog eye within a tree canopy (Villani, 2018). Another strategy to make the *identification* more robust is detecting diseases on both fruits and leaves. This is particularly essential for obtaining comprehensive insights for crop loss estimation. Early detection remains an open challenge that needs more attention which may not be addressed using RGB images alone. In such cases, leveraging images from multiple imaging sensors, such as multispectral and hyperspectral imaging can enhance the robustness of the identification process from early to advanced disease stages. Bleasdale and Whyatt (2024) attempted to utilised RGB and NIR images to classify scab disease samples both before and after symptoms appeared on the leaf. Chandel et al. (2020) utilised an unmanned aerial system to capture RGB and multispectral images to classify the healthy and powdery mildew-infected regions of apple trees in an orchard setting. In another study, the progression of fire blight infection in young apple trees was estimated through different imaging sensors, RGB, multispectral, and NIR hyperspectral images (Jarolmasjed et al., 2019). However, there remains a lack segmentation studies for other diseases in this area.

While CNN-based models have dominated the landscape of successful apple, leaf, and lesion segmentation, enhancing results by integrating different techniques or leveraging the advantages of transformer models appears to be a promising approach for achieving more accurate segmentation as discussed in Section 5. However, the constraints of complexity in vision transformers and high training time in CNNs must be alleviated, especially if the objective is to deploying segmentation

models in real-time applications. From the other side, lightweight models, which prioritise speed over precision, must strike a balance between accuracy and efficiency to provide reliable and timely disease detection. In addition, most models need proper interpretability to provide confidence to the orchardist. Interpretability can also help to find the model's weakness, leading to the incorporation of purposeful changes to the models.

CRedit authorship contribution statement

Masoumeh Keshavarzi: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Carl Mesarich:** Writing – review & editing, Validation, Supervision. **Donald Bailey:** Writing – review & editing, Validation, Methodology, Conceptualization. **Martin Johnson:** Writing – review & editing, Methodology, Conceptualization. **Gourab Sen Gupta:** Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The first author would like to acknowledge Massey University, Palmerston North, New Zealand, for providing the PhD scholarship funding that made this research possible.

Data availability

No data was used for the research described in the article.

References

- Abdalla, A., Cen, H., Wan, L., Rashid, R., Weng, H., Zhou, W., He, Y., 2019. Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. *Comput. Electron. Agric.* 167, 105091, <http://dx.doi.org/10.1016/j.compag.2019.105091>.
- Abraham, N., Khan, N.M., 2019. A novel focal tversky loss function with improved attention U-Net for lesion segmentation. In: 2019 IEEE 16th International Symposium on Biomedical Imaging. ISBI 2019, pp. 683–687, <https://doi.org/10.1109/ISBI.2019.8759329>.
- Aubhayi, B.M., Mossa, A.A., 2023. Coffee disease classification using convolutional neural network based on feature concatenation. *Informatics Med. Unlocked* 39, 101245, <http://dx.doi.org/10.1016/j.imu.2023.101245>.
- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (11), 2274–2282, <http://dx.doi.org/10.1109/TPAMI.2012.120>.
- Agarla, M., Napoletano, P., Schettini, R., 2023. Quasi real-time apple defect segmentation using deep learning. *Sensors* 23 (18), 7893, <http://dx.doi.org/10.3390/s23187893>.
- Ahmad, A., Gamal, A.E., Saraswat, D., 2023. Toward generalization of deep learning-based plant disease identification under controlled and field conditions. *IEEE Access* 11, 9042–9057, <http://dx.doi.org/10.1109/ACCESS.2023.3240100>.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (12), 2481–2495, <http://dx.doi.org/10.1109/TPAMI.2016.2644615>.
- Bakr, E., 2005. A new software for measuring leaf area, and area damaged by tetranychus urticae koch. *J. Appl. Entomol.* 129 (3), 173–175, <https://doi.org/10.1111/j.1439-0418.2005.00948.x>.
- Bansal, P., Kumar, R., Kumar, S., 2021. Disease detection in apple leaves using deep convolutional neural network. *Agriculture* 11 (7), 617, <http://dx.doi.org/10.3390/agriculture11070617>.
- Barbedo, J.G.A., 2018. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput. Electron. Agric.* 153, 46–53, <http://dx.doi.org/10.1016/j.compag.2018.08.013>.
- Barbole, D.K., Jadhav, P.M., Patil, S.B., 2022. A review on fruit detection and segmentation techniques in agricultural field. In: Second International Conference on Image Processing and Capsule Networks. pp. 269–288, http://dx.doi.org/10.1007/978-3-030-84760-9_24.
- Bargoti, S., Underwood, J.P., 2017. Image segmentation for fruit detection and yield estimation in apple orchards. *J. Field Robot.* 34 (6), 1039–1060, <https://doi.org/10.1002/rob.21699>.
- Bedi, P., Gole, P., Marwaha, S., 2024. PDSE-Lite: Lightweight framework for plant disease severity estimation based on convolutional autoencoder and few-shot learning. *Front. Plant Sci.* 14, 1319894, <http://dx.doi.org/10.3389/fpls.2023.1319894>.
- Bleasdale, A.J., Whyatt, J.D., 2024. Classifying early apple scab infections in multi-spectral imagery using convolutional neural networks. *Artif. Intell. Agric.* 15 (1), 39–51, <https://doi.org/10.1016/j.iaia.2024.10.001>.
- Bock, C.H., Barbedo, J.G.A., Del Ponte, E.M., Bohnenkamp, D., Mahlein, A.-K., 2020. From visual estimates to fully automated sensor-based measurements of plant disease severity: status and challenges for improving accuracy. *Phytopathol. Res.* 2 (1), 9, <http://dx.doi.org/10.1186/s42483-020-00049-8>.
- Bock, C.H., Chiang, K.-S., 2019. Disease incidence-severity relationships on leaflets, leaves, and fruit in the Pecan-venturia effusa pathosystem. *Plant Dis.* 103 (11), 2865–2876, <https://doi.org/10.1094/PDIS-11-18-1950-RE>.
- Bock, C.H., Chiang, K.-S., Del Ponte, E.M., 2022. Plant disease severity estimated visually: a century of research, best practices, and opportunities for improving methods and practices to maximize accuracy. *Trop. Plant Pathol.* 47 (1), 25–42, <https://doi.org/10.1007/s40858-021-00439-z>.
- Bonkra, A., Bhatt, P.K., Rosak-Szyrocka, J., Muduli, K., Pilař, L., Kaur, A., Chahal, N., Rana, A.K., 2023. Apple leaf disease detection using collaborative ML/DL and artificial intelligence methods: Scientometric analysis. *Int. J. Environ. Res. Public Heal.* 20 (4), 3222, <http://dx.doi.org/10.3390/ijerph20043222>.
- Bonkra, A., Pathak, S., Kaur, A., Shah, M.A., 2024. Exploring the trend of recognizing apple leaf disease detection through machine learning: a comprehensive analysis using bibliometric techniques. *Artif. Intell. Rev.* 57 (2), 21, <http://dx.doi.org/10.1007/s10462-023-10628-8>.
- Bowen, J.K., Mesarich, C.H., Bus, V.G., Beresford, R.M., Plummer, K.M., Templeton, M.D., 2011. *Venturia inaequalis*: the causal agent of apple scab. *Mol. Plant Pathol.* 12 (2), 105–122, <http://dx.doi.org/10.1111/j.1364-3703.2010.00656.x>.
- Bracino, A.A., Concepcion, R.S., Bedruz, R.A.R., Dadios, E.P., Vicerra, R.R.P., 2020. Development of a hybrid machine learning model for apple (*Malus domestica*) health detection and disease classification. In: 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management. HNICEM, pp. 1–6, <http://dx.doi.org/10.1109/HNICEM51456.2020.9400139>.
- Cai, M., Yi, X., Wang, G., Mo, L., Wu, P., Mwanza, C., Kapula, K.E., 2022. Image segmentation method for sweetgum leaf spots based on an improved DeeplabV3+ network. *Forests* 13 (12), 2095, <https://doi.org/10.3390/f13122095>.
- Chandel, A., Khot, L., Sallato, B., 2020. Towards rapid detection and mapping of powdery mildew in apple orchards. In: 2020 IEEE International Workshop on Metrology for Agriculture and Forestry. MetroAgriFor, pp. 288–292, <https://doi.org/10.1109/MetroAgriFor50201.2020.9277544>.
- Chatzidimopoulos, M., Lioliopoulou, F., Sotiropoulos, T., Vellios, E., 2020. Efficient control of apple scab with targeted spray applications. *Agronomy* 10 (2), 217, <http://dx.doi.org/10.3390/agronomy10020217>.
- Chen, Y., Pan, J., Wu, Q., 2023. Apple leaf disease identification via improved CycleGAN and convolutional neural network. *Soft Comput.* 27 (14), 9773–9786, <http://dx.doi.org/10.1007/s00500-023-07811-y>.
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2015. Semantic image segmentation with deep convolutional nets and fully connected CRFs. In: 3rd International Conference on Learning Representations. <http://dx.doi.org/10.48550/arXiv.1412.7062>.
- Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018a. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4), 834–848, <http://dx.doi.org/10.1109/TPAMI.2017.2699184>.
- Chen, L.-C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking atrous convolution for semantic image segmentation. *arXiv Preprint: 1706.05587*. <https://doi.org/10.48550/arXiv.1706.05587>.
- Chen, S., Zhang, K., Zhao, Y., Sun, Y., Ban, W., Chen, Y., Zhuang, H., Zhang, X., Liu, J., Yang, T., 2021. An approach for rice bacterial leaf streak disease segmentation and disease severity estimation. *Agriculture* 11 (5), 420, <http://dx.doi.org/10.3390/agriculture11050420>.
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018b. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision. ECCV, pp. 801–818, http://dx.doi.org/10.1007/978-3-030-01234-2_49.
- Chuanlei, Z., Shanwen, Z., Jucheng, Y., Yancui, S., Jia, C., 2017. Apple leaf disease identification using genetic algorithm and correlation based feature selection method. *Int. J. Agric. Biol. Eng.* 10 (2), 74–83, <http://dx.doi.org/10.3965/ijjabe.20171002.2166>.
- Dai, W., Zhu, W., Zhou, G., Liu, G., Xu, J., Zhou, H., Hu, Y., Liu, Z., Li, J., Li, L., 2024. AISOA-SSformer: An effective image segmentation method for rice leaf disease based on the transformer architecture. *Plant Phenomics* 6, 0218, <https://doi.org/10.34133/plantphenomics.0218>.

- Das, D.H., Majumder, A., Roy, S.D., Bhowmik, M.K., 2023. Segmentation and classification for bruise severity detection using infrared imaging. In: 2023 14th International Conference on Computing Communication and Networking Technologies. ICCCNT, pp. 1–7, <https://doi.org/10.1109/ICCCNT56998.2023.10306358>.
- Delalieux, S., Auwerkerken, A., Verstraeten, W.W., Somers, B., Valcke, R., Lhermitte, S., Keulemans, J., Coppin, P., 2009. Hyperspectral reflectance and fluorescence imaging to detect scab induced stress in apple leaves. *Remote. Sens.* 1 (4), 858–874, <https://doi.org/10.3390/rs1040858>.
- Deng, Y., Xi, H., Zhou, G., Chen, A., Wang, Y., Li, L., Hu, Y., 2023. An effective image-based tomato leaf disease segmentation method using MC-UNet. *Plant Phenomics* 5, 0049, <http://dx.doi.org/10.34133/plantphenomics.0049>.
- Dhiman, P., Kukreja, V., Manoharan, P., Kaur, A., Kamruzzaman, M.M., Dhaou, I.B., Iwendi, C., 2022. A novel deep learning model for detection of severity level of the disease in citrus fruits. *Electronics* 11, 495, <https://doi.org/10.3390/electronics11030495>.
- Di, J., Li, Q., 2022. A method of detecting apple leaf diseases based on improved convolutional neural network. *PLoS One* 17 (2), 1–15, <http://dx.doi.org/10.1371/journal.pone.0262629>.
- Dias, P.A., Tabb, A., Medeiros, H., 2018. Apple flower detection using deep convolutional networks. *Comput. Ind.* 99, 17–28, <https://doi.org/10.1016/j.compind.2018.03.010>.
- Ding, Y., Yang, W., Zhang, J., 2025. An improved DeepLabV3+ based approach for disease spot segmentation on apple leaves. *Comput. Electron. Agric.* 231, 110041, <https://doi.org/10.1016/j.compag.2025.110041>.
- Divyanth, L.G., Ahmad, A., Saraswat, D., 2023. A two-stage deep-learning based segmentation model for crop disease quantification based on corn field imagery. *Smart Agric. Technol.* 3, 100108, <http://dx.doi.org/10.1016/j.atech.2022.100108>.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv Preprint: 2010.11929*. <https://doi.org/10.48550/arXiv.2010.11929>.
- Douarre, C., Crispim-Junior, C.F., Gelibert, A., Tougne, L., Rousseau, D., 2019. Novel data augmentation strategies to boost supervised segmentation of plant disease. *Comput. Electron. Agric.* 165, 104967, <http://dx.doi.org/10.1016/j.compag.2019.104967>.
- Drees, L., Junker-Frohn, L.V., Kierdorf, J., Roscher, R., 2021. Temporal prediction and evaluation of Brassica growth in the field using conditional generative adversarial networks. *Comput. Electron. Agric.* 190, 106415, <http://dx.doi.org/10.1016/j.compag.2021.106415>.
- Du, J., Wang, L., Liu, Y., Zhou, Z., He, Z., Jia, Y., 2020. Brain MRI super-resolution using 3D dilated convolutional encoder–decoder network. *IEEE Access* 8, 18938–18950, <http://dx.doi.org/10.1109/ACCESS.2020.2968395>.
- Dubey, S.R., Jalal, A.S., 2012. Detection and classification of apple fruit diseases using complete local binary patterns. In: *Third International Conference on Computer and Communication Technology*. pp. 346–351, <http://dx.doi.org/10.1109/ICCCCT.2012.76>.
- Eh Teet, S., Hashim, N., 2023. Recent advances of application of optical imaging techniques for disease detection in fruits and vegetables: A review. *Food Control* 152, 109849, <https://doi.org/10.1016/j.foodcont.2023.109849>.
- El Jarroudi, M., Kouadio, A.L., Mackels, C., Tychon, B., Delfosse, P., Bock, C., 2015. A comparison between visual estimates and image analysis measurements to determine septoria leaf blotch severity in winter wheat. *Plant Pathol.* 64 (2), 355–364, <https://doi.org/10.1111/ppa.12252>.
- Elmessery, W.M., Maklakov, D.V., El-Messery, T.M., Baranenko, D.A., Gutiérrez, J., Shams, M.Y., El-Hafeez, T.A., Elsayed, S., Alhag, S.K., Moghanm, F.S., Mulyukin, M.A., Petrova, Y.Y., Elwakeel, A.E., 2024. Semantic segmentation of microbial alterations based on SegFormer. *Front. Plant Sci.* 15, 1352935, <https://doi.org/10.3389/fpls.2024.1352935>.
- Esgario, J.G.M., de Castro, P.B.C., Tassis, L.M., Krohling, R.A., 2022. An app to assist farmers in the identification of diseases and pests of coffee leaves using deep learning. *Inf. Process. Agric.* 9 (1), 38–47, <http://dx.doi.org/10.1016/j.inpa.2021.01.004>.
- Fan, Y., Beare, R., Matthews, H., Schneider, P., Kilpatrick, N., Clement, J., Claes, P., Penington, A., Adamson, C., 2019. Marker-based watershed transform method for fully automatic mandibular segmentation from CBCT images. *Dentomaxillofacial Radiol.* 48 (2), 20180261, <http://dx.doi.org/10.1259/dmfr.20180261>.
- Fan, P., Lang, G., Yan, B., Lei, X., Guo, P., Liu, Z., Yang, F., 2021. A method of segmenting apples based on gray-centered RGB color space. *Remote. Sens.* 13 (6), 1211, <http://dx.doi.org/10.3390/rs13061211>.
- Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., Zhang, C., Huang, W., 2020. Online detection of defective apples using computer vision system combined with deep learning methods. *J. Food Eng.* 286, 110102, <http://dx.doi.org/10.1016/j.jfoodeng.2020.110102>.
- Fang, S., Cui, R., Wang, Y., Zhao, Y., Yu, K., Jiang, A., 2023. Application of multiple spectral systems for the tree disease detection: A review. *Appl. Spectrosc. Rev.* 58 (2), 83–109, <https://doi.org/10.1080/05704928.2021.1930552>.
- Feng, J., Chao, X., 2022. Apple tree leaf disease segmentation dataset. *Sci. Data Bank* <http://dx.doi.org/10.11922/sciencedb.01627>.
- Feng, J., Chao, X., Zhang, Z., He, D., Zhang, J., Ye, Z., 2023. A pooling module with multidirectional and multi-scale spatial information and its application on semantic segmentation of leaf lesions. *Precis. Agric.* 24 (6), 2416–2437, <http://dx.doi.org/10.1007/s11119-023-10046-4>.
- Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* 145, 311–318, <https://doi.org/10.1016/j.compag.2018.01.009>.
- Fu, L., Li, S., Sun, Y., Mu, Y., Hu, T., Gong, H., 2022. Lightweight-convolutional neural network for apple leaf disease identification. *Front. Plant Sci.* 13, 831219, <https://doi.org/10.3389/fpls.2022.831219>.
- Gaci, B., Abdelghafour, F., Ryckewaert, M., Mas-Garcia, S., Louargant, M., Verpont, F., Laloum, Y., Moronvalle, A., Bendoula, R., Roger, J.M., 2023. Visible-Near infrared hyperspectral dataset of healthy and infected apple tree leaves images for the monitoring of apple fire blight. *Data Brief* 50, 109532, <https://doi.org/10.1016/j.dib.2023.109532>.
- Gallo, I., Rehman, A.U., Dehkordi, R.H., Landro, N., La Grassa, R., Boschetti, M., 2023. Deep object detection of crop weeds: Performance of YOLOv7 on a real case dataset from UAV images. *Remote. Sens.* 15 (2), 539, <http://dx.doi.org/10.3390/rs15020539>.
- Gao, G., Xiao, K., Jia, Y., 2020. A spraying path planning algorithm based on colour-depth fusion segmentation in peach orchards. *Comput. Electron. Agric.* 173, 105412, <https://doi.org/10.1016/j.compag.2020.105412>.
- Gené-Mola, J., Ferrer-Ferrer, M., Gregorio, E., Blok, P.M., Hemming, J., Morros, J.-R., Rosell-Polo, J.R., Vilaplana, V., Ruiz-Hidalgo, J., 2023. Looking behind occlusions: A study on amodal segmentation for robust on-tree apple fruit size estimation. *Comput. Electron. Agric.* 209, 107854, <https://doi.org/10.1016/j.compag.2023.107854>.
- Goëau, H., Bonnet, P., Joly, A., 2023. Overview of PlantCLEF 2023: image-based plant identification at global scale. In: *24th Working Notes of the Conference and Labs of the Evaluation Forum, CLEF-WN 2023*. Vol. 3497, pp. 1972–1981.
- Gonçalves, J.P., Pinto, F.A.C., Queiroz, D.M., Villar, F.M.M., Barbedo, J.G.A., Del Ponte, E.M., 2021. Deep learning architectures for semantic segmentation and automatic estimation of severity of foliar symptoms caused by diseases or pests. *Biosyst. Eng.* 210, 129–142, <http://dx.doi.org/10.1016/j.biosystemseng.2021.08.011>.
- Grove, G.G., Eastwell, K.C., Jones, A.L., Sutton, T.B., 2003. Diseases of apple. In: *Apples: Botany, Production and Uses*. pp. 459–488, <http://dx.doi.org/10.1079/9780851995922.0459>.
- Gupta, S.K., Yadav, S.K., Soni, S.K., Shanker, U., Singh, P.K., 2023. Multiclass weed identification using semantic segmentation: An automated approach for precision agriculture. *Ecol. Inform.* 78, 102366, <https://doi.org/10.1016/j.ecoinf.2023.102366>.
- Hamada, N., Moreira, R., Nesi, C., Mio, L.M., 2019. Pathogen dispersal and Glomerella leaf spot progress within apple canopy in Brazil. *Plant Dis.* 103, <https://doi.org/10.1094/PDIS-08-18-1375-RE>.
- Häni, N., Roy, P., Isler, V., 2019. A comparative study of fruit detection and counting methods for yield mapping in apple orchards. *J. Field Robot.* 37, 263–282, <https://doi.org/10.1002/rob.21902>.
- Häni, N., Roy, P., Isler, V., 2020. MinneApple: A benchmark dataset for apple detection and segmentation. *IEEE Robot. Autom. Lett.* 5 (2), 852–858, <https://doi.org/10.1109/LRA.2020.2965061>.
- Hasan, S., Jahan, S., Islam, M.I., 2022b. Disease detection of apple leaf with combination of color segmentation and modified DWT. *J. King Saud Univ. - Comput. Inf. Sci.* 34 (9), 7212–7224, <http://dx.doi.org/10.1016/j.jksuci.2022.07.004>.
- Hasan, R.I., Yusuf, S.M., Rahim, M.S.M., Alzubaidi, L., 2022a. Automated masks generation for coffee and apple leaf infected with single or multiple diseases-based color analysis approaches. *Inform. Med. Unlocked* 28, 100837, <http://dx.doi.org/10.1016/j.imu.2021.100837>.
- Hayit, T., Erbay, H., Varçın, F., Hayit, F., Akci, N., 2021. Determination of the severity level of yellow rust disease in wheat by using convolutional neural networks. *J. Plant Pathol.* 103 (3), 923–934, <https://doi.org/10.1007/s42161-021-00886-2>.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition*. CVPR, pp. 770–778, <http://dx.doi.org/10.1109/CVPR.2016.90>.
- Hou, J., Yang, C., He, Y., Hou, B., 2023. Detecting diseases in apple tree leaves using FPN-ISResNet-Faster RCNN. *Eur. J. Remote. Sens.* 56 (1), 2186955, <http://dx.doi.org/10.1080/22797254.2023.2186955>.
- Hou, Q., Zhang, L., Cheng, M.M., Feng, J., 2020. Strip pooling: Rethinking spatial pooling for scene parsing. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. CVPR, pp. 4002–4011, <http://dx.doi.org/10.1109/CVPR42600.2020.00406>.
- Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E., 2020. Squeeze-and-excitation networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 42 (8), 2011–2023, <http://dx.doi.org/10.1109/TPAMI.2019.2913372>.
- Hu, G., Wan, M., Wei, K., Ye, R., 2023. Computer vision based method for severity estimation of tea leaf blight in natural scene images. *Eur. J. Agron.* 144, 126756, <http://dx.doi.org/10.1016/j.eja.2023.126756>.
- Huang, Q., Wang, X., Gao, Q., Carraro, A., Sozzi, M., Marinello, F., 2024b. Indicators to digitization footprint and how to get digitization footprint (Part 2). *Comput. Electron. Agric.* 224, 109206, <https://doi.org/10.1016/j.compag.2024.109206>.

- Huang, H., Wang, Z., Zhao, G., 2024a. Classification and segmentation of apple diseases based on deep learning. In: Proceedings of the 2024 16th International Conference on Machine Learning and Computing. pp. 433–440, <https://doi.org/10.1145/3651671.3651715>.
- Hughes, D., Salathé, M., et al., 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv Preprint: 1511.08060. <https://doi.org/10.48550/arXiv.1511.08060>.
- Jamadar, R.A., Sharma, A., 2023. Semantic segmentation based leaf disease severity estimation using deep learning algorithms. In: 2023 International Conference on Emerging Smart Computing and Informatics. ESCI, pp. 1–5, <http://dx.doi.org/10.1109/ESCI56872.2023.10099491>.
- Jarolmasjed, S., Sankaran, S., Marzougui, A., Kostick, S., Si, Y., Quirós Vargas, J.J., Evans, K., 2019. High-throughput phenotyping of fire blight disease symptoms using sensing techniques in apple. *Front. Plant Sci.* 10, 576, <https://doi.org/10.3389/fpls.2019.00576>.
- Javidan, S.M., Banakar, A., Vakilian, K.A., Ampatzidis, Y., 2023. Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning. *Smart Agric. Technol.* 3, 100081, <http://dx.doi.org/10.1016/j.atech.2022.100081>.
- Ji, M., Zhang, K., Wu, Q., Deng, Z., 2020. Multi-label learning for crop leaf diseases recognition and severity estimation based on convolutional neural networks. *Soft Comput.* 24 (20), 15327–15340, <https://doi.org/10.1007/s00500-020-04866-z>.
- Jia, W., Tian, Y., Luo, R., Zhang, Z., Lian, J., Zheng, Y., 2020. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. *Comput. Electron. Agric.* 172, 105380, <https://doi.org/10.1016/j.compag.2020.105380>.
- Jiang, D., Chang, Q., Zhang, Z., Liu, Y., Zhang, Y., Zheng, Z., 2023. Monitoring the degree of mosaic disease in apple leaves using hyperspectral images. *Remote. Sens.* 15 (10), 2504, <https://doi.org/10.3390/rs15102504>.
- Jiang, D., Sun, M., Li, S., Yang, Z., Cao, L., 2024. EFS-Former: An efficient network for fruit tree leaf disease segmentation and severity assessment. *Agronomy* 14 (9), 1992, <https://doi.org/10.3390/agronomy14091992>.
- Jourlin, M., Breugnot, J., Abdallah, B., Corvo, J., Couka, E., Carré, M., 2013. Chapter one - image segmentation in the field of the logarithmic image processing model: Special focus on the hierarchical ascend classification techniques. In: *Advances in Imaging and Electron Physics*. Vol. 177, pp. 1–44, <http://dx.doi.org/10.1016/B978-0-12-407702-7.00001-2>.
- Kang, H., Chen, C., 2019. Fruit detection and segmentation for apple harvesting using visual sensor in orchards. *Sensors* 19 (20), 4599, <https://doi.org/10.3390/s19204599>.
- Keshavarzi, M., Mesarich, C., Johnson, M., Gupta, G.S., Bailey, D., 2024. Leaf or leaves? A data-centric approach looking at ground truth effect on apple disease segmentation. In: 2024 39th International Conference on Image and Vision Computing New Zealand. IVCNZ, pp. 1–6, <https://doi.org/10.1109/IVCNZ64857.2024.10794466>.
- Khan, M.A., Akram, T., Sharif, M., Awais, M., Javed, K., Ali, H., Saba, T., 2018. CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coefficient and deep CNN features. *Comput. Electron. Agric.* 155, 220–236, <http://dx.doi.org/10.1016/j.compag.2018.10.013>.
- Khan, K., Khan, R.U., Albattah, W., Qamar, A.M., 2022b. End-to-end semantic leaf segmentation framework for plants disease classification. *Complexity* 2022, 1168700, <http://dx.doi.org/10.1155/2022/1168700>.
- Khan, S., Narvekar, M., 2022. Novel fusion of color balancing and superpixel based approach for detection of tomato plant diseases in natural complex environment. *J. King Saud Univ. - Comput. Inf. Sci.* 34 (6, Part B), 3506–3516, <http://dx.doi.org/10.1016/j.jksuci.2020.09.006>.
- Khan, A.I., Quadri, S.M.K., Banday, S., Latief Shah, J., 2022a. Deep diagnosis: A real-time apple leaf disease detection system based on deep learning. *Comput. Electron. Agric.* 198, 107093, <http://dx.doi.org/10.1016/j.compag.2022.107093>.
- Kim, M., 2021. Apple leaf disease classification using superpixel and CNN. In: *Advances in Computer Vision and Computational Biology*. pp. 99–106, http://dx.doi.org/10.1007/978-3-030-71051-4_8.
- Kim, J., Seol, J., Lee, S., Hong, S.W., Son, H.I., 2020. An intelligentspraying system with deep learning-based semantic segmentation of fruit trees in orchards. In: 2020 IEEE International Conference on Robotics and Automation. ICRA, pp. 3923–3929, <http://dx.doi.org/10.1109/ICRA40945.2020.9197556>.
- Kleynen, O., Leemans, V., Destain, M.F., 2005. Development of a multi-spectral vision system for the detection of defects on apples. *J. Food Eng.* 69 (1), 41–49, <https://doi.org/10.1016/j.jfoodeng.2004.07.008>.
- Kodors, S., Laci, G., Sokolova, O., Zhukov, V., Apeinans, I., Bartulsons, T., 2021. Apple scab detection using CNN and transfer learning. *Agron. Res.* 19, 507–519, <http://dx.doi.org/10.15159/AR.21.045>.
- Kotwal, J., Kashyap, R., Pathan, S., 2023. Agricultural plant diseases identification: From traditional approach to deep learning. *Mater. Today: Proc.* 80, 344–356, <http://dx.doi.org/10.1016/j.matpr.2023.02.370>.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60 (6), 84–90, <http://dx.doi.org/10.1145/3065386>.
- Lamari, L., 2002. Assess: Image analysis software for plant disease quantification. *Am. Phytopathol. Soc.*
- Larijani, M.R., Asli-Ardeh, E.A., Kozegar, E., Loni, R., 2019. Evaluation of image processing technique in identifying rice blast disease in field conditions based on KNN algorithm improvement by K-means. *Food Science Nutr.* 7 (12), 3922–3930, <http://dx.doi.org/10.1002/fsn3.1251>.
- Lee, S.H., Goëau, H., Bonnet, P., Joly, A., 2020. New perspectives on plant disease characterization based on deep learning. *Comput. Electron. Agric.* 170, 105220, <http://dx.doi.org/10.1016/j.compag.2020.105220>.
- Li, Q., Jia, W., Sun, M., Hou, S., Zheng, Y., 2021. A novel green apple segmentation algorithm based on ensemble U-Net under complex orchard environment. *Comput. Electron. Agric.* 180, 105900, <http://dx.doi.org/10.1016/j.compag.2020.105900>.
- Li, J., Luo, W., Wang, Z., Fan, S., 2019. Early detection of decay on apples using hyperspectral reflectance imaging combining both principal component analysis and improved watershed segmentation method. *Postharvest Biol. Technol.* 149, 235–246, <http://dx.doi.org/10.1016/j.postharvbio.2018.12.007>.
- Li, L., Wang, B., Li, Y., Yang, H., 2023a. Diagnosis and mobile application of apple leaf disease degree based on a small-sample dataset. *Plants* 12 (4), 786, <http://dx.doi.org/10.3390/plants12040786>.
- Li, K., Zhang, L., Li, B., Li, S., Ma, J., 2022a. Attention-optimized DeepLab V3+ for automatic estimation of cucumber disease severity. *Plant Methods* 18 (1), 109, <https://doi.org/10.1186/s13007-022-00941-8>.
- Li, L., Zhang, S., Wang, B., 2022b. Apple leaf disease identification with a small and imbalanced dataset based on lightweight convolutional networks. *Sensor* 22 (1), 173, <http://dx.doi.org/10.3390/s22010173>.
- Li, X., Zhou, H., Chang, Y., Zhang, C., 2023b. A method of apple leaf spot segmentation based on improved DeepLabV3+. In: *International Conference on Cloud Computing, Performance Computing, and Deep Learning. CCPCDL 2023*, Vol. 12712, 1271203, <http://dx.doi.org/10.1117/12.2679129>.
- Lian, S., Guan, L., Pei, J., Zeng, G., Li, M., 2024. Identification of apple leaf diseases using C-Grabcut algorithm and improved transfer learning base on low shot learning. *Multimedia Tools Appl.* 83 (9), 27411–27433, <http://dx.doi.org/10.1007/s11042-023-16602-4>.
- Liang, X., Gao, S., 2024. Segmentation of aapple diseased leaves by multiscale feature extraction UNet modeling. In: *Proceedings of the 2024 3rd International Conference on Cyber Security, Artificial Intelligence and Digital Economy*. pp. 426–431, <https://doi.org/10.1145/3672919.3672997>.
- Liang, X., Jia, X., Huang, W., He, X., Li, L., Fan, S., Li, J., Zhao, C., Zhang, C., 2022. Real-time grading of defect apples using semantic segmentation combination with a pruned YOLO V4 network. *Foods* 11 (19), 3150, <https://doi.org/10.3390/foods11193150>.
- Liang, Q., Xiang, S., Hu, Y., Coppola, G., Zhang, D., Sun, W., 2019. PD²SE-Net: Computer-assisted plant disease diagnosis and severity estimation network. *Comput. Electron. Agric.* 157, 518–529, <https://doi.org/10.1016/j.compag.2019.01.034>.
- Lin, T., Goyal, P., Girshick, R., He, K., Dollar, P., 2017. Focal loss for dense object detection. In: 2017 IEEE International Conference on Computer Vision. ICCV, pp. 2999–3007, <https://doi.org/10.1109/ICCV.2017.324>.
- Liu, B.-Y., Fan, K.-J., Su, W.-H., Peng, Y., 2022a. Two-stage convolutional neural networks for diagnosing the severity of alternaria leaf blotch disease of the apple tree. *Remote. Sens.* 14 (11), 2519, <http://dx.doi.org/10.3390/rs14112519>.
- Liu, B., Fan, H., Zhang, Y., Cai, J., Cheng, H., 2024a. Deep learning architectures for diagnosing the severity of apple frog-eye leaf spot disease in complex backgrounds. *Front. Plant Sci.* 14, 1289497, <https://doi.org/10.3389/fpls.2023.1289497>.
- Liu, E., Gold, K.M., Combs, D., Cadle-Davidson, L., Jiang, Y., 2022b. Deep semantic segmentation for the quantification of grape foliar diseases in the vineyard. *Front. Plant Sci.* 13, 978761, <http://dx.doi.org/10.3389/fpls.2022.978761>.
- Liu, J., Yu, S., Liu, X., Lu, G., Xin, Z., Yuan, J., 2024b. Super-resolution semantic segmentation of droplet deposition image for low-cost spraying measurement. *Agriculture* 14 (1), 106, <https://doi.org/10.3390/agriculture14010106>.
- Liu, B., Zhang, Y., He, D., Li, Y., 2018. Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry* 10 (1), 11, <http://dx.doi.org/10.3390/sym10010011>.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 3431–3440, <http://dx.doi.org/10.1109/CVPR.2015.7298965>.
- Lu, Y., Lu, R., 2017. Non-destructive defect detection of apples by spectroscopic and imaging technologies: a review. *Trans. the ASABE (Am. Soc. Agric. Biol. Eng.)* 60 (5), 1765–1790, <http://dx.doi.org/10.13031/trans.12431>.
- Lu, J., Lu, B., Ma, W., Sun, Y., 2024. EAIIS-Former: An efficient and accurate image segmentation method for fruit leaf diseases. *Comput. Electron. Agric.* 218, 108739, <http://dx.doi.org/10.1016/j.compag.2024.108739>.
- Lu, B., Lu, J., Xu, X., Jin, Y., 2023. MixSeg: A lightweight and accurate mix structure network for semantic segmentation of apple leaf disease in complex environments. *Front. Plant Sci.* 14, 1233241, <http://dx.doi.org/10.3389/fpls.2023.1233241>.
- Luo, Y., Sun, J., Shen, J., Wu, X., Wang, L., Zhu, W., 2021. Apple leaf disease recognition and sub-class categorization based on improved multi-scale feature fusion network. *IEEE Access* 9, 95517–95527, <http://dx.doi.org/10.1109/ACCESS.2021.3094802>.
- Luo, W., Zhang, H., Liu, X., 2019. Hyperspectral/multispectral reflectance imaging combining with watershed segmentation algorithm for detection of early bruises on apples with different peel colors. *Food Anal. Methods* 12 (5), 1218–1228, <http://dx.doi.org/10.1007/s12161-019-01456-0>.

- Ma, J., Dai, Y., Tan, Y.-P., 2019. Atrous convolutions spatial pyramid network for crowd counting and density estimation. *Neurocomputing* 350, 91–101, <http://dx.doi.org/10.1016/j.neucom.2019.03.065>.
- Ma, Y.-d., Liu, Q., Qian, Z.-b., 2004. Automated image segmentation using improved PCNN model based on cross-entropy. In: *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*. pp. 743–746, <https://doi.org/10.1109/ISIMP.2004.1434171>.
- Mahmud, M.S., He, L., Zahid, A., Heinemann, P., Choi, D., Krawczyk, G., Zhu, H., 2023. Detection and infected area segmentation of apple fire blight using image processing and deep transfer learning for site-specific management. *Comput. Electron. Agric.* 209, 107862, <http://dx.doi.org/10.1016/j.compag.2023.107862>.
- Maß, V., Alirezazadeh, P., Seidl-Schulz, J., Leipnitz, M., Fritzsche, E., Ibraheem, R.A.A., Geyer, M., Pflanz, M., Reim, S., 2024. Annotated image dataset of fire blight symptoms for object detection in orchards. *Data Brief* 56, 110826, <https://doi.org/10.1016/j.dib.2024.110826>.
- Milioto, A., Lottes, P., Stachniss, C., 2018. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs. In: *2018 IEEE International Conference on Robotics and Automation. ICRA*, pp. 2229–2235.
- Moazzam, S.I., Khan, U.S., Qureshi, W.S., Nawaz, T., Kunwar, F., 2023. Towards automated weed detection through two-stage semantic segmentation of tobacco and weed pixels in aerial imagery. *Smart Agric. Technol.* 4, 100142, <https://doi.org/10.1016/j.atech.2022.100142>.
- Mzoughi, O., Yahiaoui, I., 2023. Deep learning-based segmentation for disease identification. *Ecol. Inform.* 75, 102000, <http://dx.doi.org/10.1016/j.ecoinf.2023.102000>.
- Nachtigall, L.G., Araujo, R., Nachtigall, G., 2020. Use of images of leaves and fruits of apple trees for automatic identification of symptoms of diseases and nutritional disorders. In: *Deep Learning and Neural Networks*. pp. 2980–2988, <http://dx.doi.org/10.4018/978-1-7998-0414-7.ch060>.
- Ngugi, L.C., Abdelwahab, M., Abo-Zahhad, M., 2020. Tomato leaf segmentation algorithms for mobile phone applications using deep learning. *Comput. Electron. Agric.* 178, 105788, <http://dx.doi.org/10.1016/j.compag.2020.105788>.
- Nouri, M., Gorretta, N., Vaysse, P., Giraud, M., Germain, C., Keresztes, B., Roger, J.-M., 2018. Near infrared hyperspectral dataset of healthy and infected apple tree leaves images for the early detection of apple scab disease. *Data Brief* 16, 967–971, <https://doi.org/10.1016/j.dib.2017.12.043>.
- Oktaç, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., 2018. Attention U-Net: Learning where to look for the pancreas. *arXiv Preprint: 1804.03999*. <https://doi.org/10.48550/arXiv.1804.03999>.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 9 (1), 62–66, <http://dx.doi.org/10.1109/TSMC.1979.4310076>.
- Pal, A., Kumar, V., 2023. AgriDet: Plant leaf disease severity classification using agriculture detection framework. *Eng. Appl. Artif. Intell.* 119, 105754, <http://dx.doi.org/10.1016/j.engappai.2022.105754>.
- Parashar, N., Johri, P., 2024. Enhancing apple leaf disease detection: A CNN-based model integrated with image segmentation techniques for precision agriculture. *Int. J. Math. Eng. Manag. Sci.* 9 (4), 943–964, <https://doi.org/10.33889/IJMMS.2024.9.4.050>.
- Pethybridge, S.J., Nelson, S.C., 2015. Leaf doctor: A new portable application for quantifying plant disease severity. *Plant Dis.* 99 (10), 1310–1316, <https://doi.org/10.1094/pdis-03-15-0319-re>.
- Prabhakar, M., Purushothaman, R., Awasthi, D.P., 2020. Deep learning based assessment of disease severity for early blight in tomato crop. *Multimedia Tools Appl.* 79 (39), 28773–28784, <http://dx.doi.org/10.1007/s11042-020-09461-w>.
- Rai, C.K., Pahuja, R., 2023. Detection and segmentation of rice diseases using deep convolutional neural networks. *SN Comput. Sci.* 4 (5), 499, <http://dx.doi.org/10.1007/s42979-023-02014-6>.
- Rayhana, R., Ma, Z., Liu, Z., Xiao, G.G., Ruan, Y., Sangha, J.S., 2023. A review on plant disease detection using hyperspectral imaging. *IEEE Trans. AgriFood Electron.* 1, 108–134, <https://doi.org/10.1109/TAFE.2023.3329849>.
- Rehman, Z.U., Attique Khan, M., Fawad, A., Damaševičius, R., Naqvi, S.R., Nisar, W., Javed, K., 2021. Recognizing apple leaf diseases using a novel parallel real-time processing framework based on MASK-RCNN and transfer learning: An application for smart agriculture. *IET Image Process.* 15, 2157–2168, <http://dx.doi.org/10.1049/ipr2.12183>.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer, pp. 234–241, http://dx.doi.org/10.1007/978-3-319-24574-4_28.
- Rother, C., Kolmogorov, V., Blake, A., 2004. GrabCut: Interactive foreground extraction using iterated graph cuts. In: *Special Interest Group on Computer Graphics and Interactive Techniques*. Vol. 23, pp. 309–314, <http://dx.doi.org/10.1145/1186562.1015720>.
- Rousseau, C., Belin, E., Bove, E., Rousseau, D., Fabre, F., Berruyer, R., Guillaumès, J., Manceau, C., Jacques, M.-A., Boureau, T., 2013. High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis. *Plant Methods* 9 (1), 17, <https://doi.org/10.1186/1746-4811-9-17>.
- Roy, K., Chaudhuri, S.S., Pramanik, S., 2021. Deep learning based real-time industrial framework for rotten and fresh fruit detection using semantic segmentation. *Microsyst. Technol.* 27 (9), 3365–3375, <https://doi.org/10.1007/s00542-020-05123-x>.
- Saleem, M.H., Potgieter, J., Arif, K.M., 2022. A performance-optimized deep learning-based plant disease detection approach for horticultural crops of New Zealand. *IEEE Access* 10, 89798–89822, <http://dx.doi.org/10.1109/ACCESS.2022.3201104>.
- Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9 (7), 671–675, <https://doi.org/10.1038/nmeth.2089>.
- Sehar, U., Naseem, M.L., 2022. How deep learning is empowering semantic segmentation. *Multimedia Tools Appl.* 81 (21), 30519–30544, <http://dx.doi.org/10.1007/s11042-022-12821-3>.
- Selvaraj, G., Farzan, S., 2024. Semantic segmentation for robotic apple harvesting: A deep learning approach leveraging U-Net, synthetic data, and domain adaptation. In: *21st International Conference on Ubiquitous Robots*. UR, pp. 611–618, <https://doi.org/10.1109/UR61395.2024.10597519>.
- Shahoveisi, F., Taheri Gorji, H., Shahabi, S., Hosseinirad, S., Markell, S., Vasefi, F., 2023. Application of image processing and transfer learning for the detection of rust disease. *Sci. Rep.* 13 (1), 5133, <http://dx.doi.org/10.1038/s41598-023-31942-9>.
- Sharma, H., Padha, D., Bashir, N., 2023. D-KAP: A deep learning-based Kashmiri apple plant disease prediction framework. In: *2022 Seventh International Conference on Parallel, Distributed and Grid Computing. PDGC*, pp. 576–581, <http://dx.doi.org/10.1109/PDGC56933.2022.10053334>.
- Sheng, X., Kang, C., Zheng, J., Lyu, C., 2023. An edge-guided method to fruit segmentation in complex environments. *Comput. Electron. Agric.* 208, 107788, <http://dx.doi.org/10.1016/j.compag.2023.107788>.
- Shi, T., Liu, Y., Zheng, X., Hu, K., Huang, H., Liu, H., Huang, H., 2023. Recent advances in plant disease severity assessment using convolutional neural networks. *Sci. Rep.* 13 (1), 2336, <https://doi.org/10.1038/s41598-023-29230-7>.
- Shoabi, M., Hussain, T., Shah, B., Ullah, I., Shah, S.M., Ali, F., Park, S.H., 2022. Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease. *Front. Plant Sci.* 13, 1031748, <https://doi.org/10.3389/fpls.2022.1031748>.
- Siddiqua, A., Kabir, M.A., Ferdous, T., Ali, I.B., Weston, L.A., 2022. Evaluating plant disease detection mobile applications: quality and limitations. *Agronomy* 12 (8), 1869, <https://doi.org/10.3390/agronomy12081869>.
- Singh, S., Bergerman, M., Hoheisel, G., Lewis, K., Baugher, T., 2011. CASC IFW dataset. <https://engineering.purdue.edu/RVL/Database/IFW/index.html>.
- Singh, S., Gupta, S., 2018. A novel algorithm for segmentation of diseased apple leaf images. *J. Adv. Res. Dyn. Control. Syst.* 10 (6 Special Issue), 2248–2259.
- Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S., Batra, N., 2020. PlantDoc: A dataset for visual plant disease detection. In: *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*. pp. 249–253, <http://dx.doi.org/10.1145/3371158.3371196>.
- Skoneczny, H., Kubiak, K., Spiralski, M., Kotlarz, J., Mikiciński, A., Puławska, J., 2020. Fire blight disease detection for apple trees: Hyperspectral analysis of healthy, infected and dry leaves. *Remote. Sens.* 12 (13), 2101, <https://doi.org/10.3390/rs12132101>.
- Sodjinoou, S.G., Mohammadi, V., Sanda Mahama, A.T., Gouton, P., 2022. A deep semantic segmentation-based algorithm to segment crops and weeds in agronomic color images. *Inf. Process. Agric.* 9 (3), 355–364, <http://dx.doi.org/10.1016/j.inpa.2021.08.003>.
- Son, C.-H., 2021. Leaf spot attention networks based on spot feature encoding for leaf disease identification and detection. *Appl. Sci.* 11 (17), 7960, <http://dx.doi.org/10.3390/app11177960>.
- Stasenko, N., Chernova, E., Shadrin, D., Ovchinnikov, G., Krivolapov, I., Pukalchik, M., 2021. Deep learning for improving the storage process: Accurate and automatic segmentation of spoiled areas on apples. In: *2021 IEEE International Instrumentation and Measurement Technology Conference. I2MTC*, pp. 1–6, <http://dx.doi.org/10.1109/I2MTC50364.2021.9460071>.
- Storey, G., Meng, Q., Li, B., 2022. Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture. *Sustainability* 14 (3), 1458, <https://doi.org/10.3390/su14031458>.
- Su, W.-H., Zhang, J., Yang, C., Page, R., Szinyei, T., Hirsch, C., Steffenson, B., 2021. Automatic evaluation of wheat resistance to Fusarium head blight using dual Mask-RCNN deep learning frameworks in computer vision. *Remote. Sens.* 13, 1–21, <http://dx.doi.org/10.3390/rs13010026>.
- Sun, L., He, J., Zhang, L., 2024. CASP-MNet: multi-scale network with cross attention mechanism and spatial dimension feature fusion for maize leaf disease detection. *Crop. Prot.* 180, 106667, <http://dx.doi.org/10.1016/j.cropro.2024.106667>.
- Sun, H., Wei, J., Zhang, J., Yang, W., 2014. A comparison of disease severity measurements using image analysis and visual estimates using a category scale for genetic analysis of resistance to bacterial spot in tomato. *Eur. J. Plant Pathol.* 139 (1), 125–136, <https://doi.org/10.1007/s10658-013-0371-8>.
- Tang, L., Yi, J., Li, X., 2023. Improved multi-scale inverse bottleneck residual network based on triplet parallel attention for apple leaf disease identification. *J. Integr. Agric.* 23, 901–922, <http://dx.doi.org/10.1016/j.jia.2023.06.023>.
- Tassiss, L.M., Tozzi de Souza, J.E., Krohling, R.A., 2021. A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images. *Comput. Electron. Agric.* 186, 106191, <http://dx.doi.org/10.1016/j.compag.2021.106191>.

- Thapa, R., Zhang, K., Snaveley, N., Belongie, S., Khan, A., 2020. The plant pathology challenge 2020 data set to classify foliar disease of apples. *Appl. Plant Sci.* 8 (9), e11390, <https://doi.org/10.1002/aps3.11390>.
- Thapa, R., Zhang, K., Snaveley, N., Belongie, S., Khan, A., 2021. Plant pathology 2021 - FGVCS. <https://kaggle.com/competitions/plant-pathology-2021-fgvc8>.
- Tian, Y., Li, E., Liang, Z., Tan, M., He, X., 2021. Diagnosis of typical apple diseases: A deep learning method based on multi-scale dense classification network. *Front. Plant Sci.* 12, 698474, <http://dx.doi.org/10.3389/fpls.2021.698474>.
- Tian, X., Liu, X., He, X., Zhang, C., Li, J., Huang, W., 2023. Detection of early bruises on apples using hyperspectral reflectance imaging coupled with optimal wavelengths selection and improved watershed segmentation algorithm. *J. Sci. Food Agric.* 103 (13), 6689–6705, <http://dx.doi.org/10.1002/jsfa.12764>.
- Tian, Y., Yang, G., Wang, Z., Li, E., Liang, Z., 2019. Detection of apple lesions in orchards based on deep learning methods of CycleGAN and YOLOV3-Dense. *J. Sensors* 2019, 7630926, <http://dx.doi.org/10.1155/2019/7630926>.
- Tian, M., Zhang, J., Yang, Z., Li, M., Li, J., Zhao, L., 2024. Detection of early bruises on apples using near-infrared camera imaging technology combined with adaptive threshold segmentation algorithm. *J. Food Process. Eng.* 47 (1), e14500, <https://doi.org/10.1111/jfpe.14500>.
- Tong, S., Zhang, J., Li, W., Wang, Y., Kang, F., 2023. An image-based system for locating pruning points in apple trees using instance segmentation and RGB-D images. *Biosyst. Eng.* 236, 277–286, <https://doi.org/10.1016/j.biosystemseng.2023.11.006>.
- Upadhyay, S.K., Kumar, A., 2022. A novel approach for rice plant diseases classification with deep convolutional neural network. *Int. J. Inf. Technol.* 14 (1), 185–199, <http://dx.doi.org/10.1007/s41870-021-00817-5>.
- Uryasheva, A., Kalashnikova, A., Shadrin, D., Evteeva, K., Moskovtsev, E., Rodichenko, N., 2022. Computer vision-based platform for apple leaves segmentation in field conditions to support digital phenotyping. *Comput. Electron. Agric.* 201, 107269, <https://doi.org/10.1016/j.compag.2022.107269>.
- Villani, S., 2018. Leaf spots of bother: The troublesome 3. <https://apples.ces.ncsu.edu/2018/03/leaf-spots-of-bother-the-troublesome-3/>. [Accessed November 2023].
- Wang, C., Du, P., Wu, H., Li, J., Zhao, C., Zhu, H., 2021. A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net. *Comput. Electron. Agric.* 189, 106373, <http://dx.doi.org/10.1016/j.compag.2021.106373>.
- Wang, D., He, D., 2022. Fusion of Mask RCNN and attention mechanism for instance segmentation of apples under complex background. *Comput. Electron. Agric.* 196, 106864, <https://doi.org/10.1016/j.compag.2022.106864>.
- Wang, J., Jia, J., Zhang, Y., Wang, H., Zhu, S., 2024. RAAWC-UNet: an apple leaf and disease segmentation method based on residual attention and atrous spatial pyramid pooling improved UNet with weight compression loss. *Front. Plant Sci.* 15, 1305358, <http://dx.doi.org/10.3389/fpls.2024.1305358>.
- Wang, Z., Lv, Y., Wu, R., Zhang, Y., 2023. Review of GrabCut in image processing. *Math. Methods Pattern Recognit.* 11 (8), 1965, <http://dx.doi.org/10.3390/math11081965>.
- Wang, G., Sun, Y., Wang, J., 2017. Automatic image-based plant disease severity estimation using deep learning. *Comput. Intell. Neurosci.* 2017, 2917536, <http://dx.doi.org/10.1155/2017/2917536>.
- Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q., 2020. ECA-Net: Efficient channel attention for deep convolutional neural networks. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR*, pp. 11531–11539, <http://dx.doi.org/10.1109/CVPR42600.2020.01155>.
- Woo, S., Park, J., Lee, J.-Y., Kweon, I.S., 2018. CBAM: Convolutional block attention module. In: *Computer Vision – European Conference on Computer Vision. ECCV*, pp. 3–19, https://doi.org/10.1007/978-3-030-01234-2_1.
- Xiao, D., Pan, Y., Feng, J., Yin, J., Liu, Y., He, L., 2022. Remote sensing detection algorithm for apple fire blight based on UAV multispectral image. *Comput. Electron. Agric.* 199, 107137, <https://doi.org/10.1016/j.compag.2022.107137>.
- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J.M., Luo, P., 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Adv. Neural Inf. Process. Syst.* 34, 12077–12090.
- Xing, B., Wang, D., Yin, T., 2023. The evaluation of the grade of leaf disease in apple trees based on PCA-logistic regression analysis. *Forests* 14 (7), 1290, <https://doi.org/10.3390/f14071290>.
- Xu, L., Li, Y., Xu, J., Guo, L., 2020. Two-level attention and score consistency network for plant segmentation. *Comput. Electron. Agric.* 170, 105281, <http://dx.doi.org/10.1016/j.compag.2020.105281>.
- Xu, C., Yu, C., Zhang, S., 2022. Lightweight multi-scale dilated U-Net for crop disease leaf image segmentation. *Electronics* 11 (23), 3947, <http://dx.doi.org/10.3390/electronics11233947>.
- Xue, W., Haynes, K.G., Qu, X., 2019. Characterization of early blight resistance in potato cultivars. *Plant Dis.* 103 (4), 629–637, <https://doi.org/10.1094/PDIS-05-18-0794-RE>.
- Xue, X., Luo, Q., Bu, M., Li, Z., Lyu, S., Song, S., 2023. Citrus tree canopy segmentation of orchard spraying robot based on RGB-D image and the improved DeepLabv3+. *Agronomy* 13 (8), 2059, <https://doi.org/10.3390/agronomy13082059>.
- Yadav, A., Thakur, U., Saxena, R., Pal, V., Bhateja, V., Lin, J.C.-W., 2022. AFD-Net: Apple foliar disease multi-classification using deep learning on plant pathology dataset. *Plant Soil* 477 (1), 595–611, <http://dx.doi.org/10.1007/s11104-022-05407-3>.
- Yang, W., Hu, W., Xie, L., Yang, Z., 2022. Multi-scale and multi-match for few-shot plant disease image semantic segmentation. *Agronomy* 12, 2847, <http://dx.doi.org/10.3390/agronomy12112847>.
- Yang, B., Wang, Z., Guo, J., Guo, L., Liang, Q., Zeng, Q., Zhao, R., Wang, J., Li, C., 2023. Identifying plant disease and severity from leaves: A deep multitask learning framework using triple-branch Swin Transformer and deep supervision. *Comput. Electron. Agric.* 209, 107809, <http://dx.doi.org/10.1016/j.compag.2023.107809>.
- Yang, T., Wang, Y., Lian, J., 2024. Plant diseased lesion image segmentation and recognition based on improved Multi-Scale Attention Net. *Agric. Sci. Technol.* 14, 1716, <http://dx.doi.org/10.3390/app14051716>.
- Yatoo, A.A., Sharma, A., 2024. An indigenous dataset for the detection and classification of apple leaf diseases. *Data Brief* 53, 110165, <http://dx.doi.org/10.1016/j.dib.2024.110165>.
- Yogesh, Singhal, P., Dubey, A.K., Goyal, A., 2017. A comparative approach for image segmentation to identify the defected portion of apple. In: *6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions). ICRITO*, pp. 604–608, <http://dx.doi.org/10.1109/ICRITO.2017.8342499>.
- Yu, C., Gao, C., Wang, J., Yu, G., Shen, C., Sang, N., 2021. BiSeNet V2: Bilateral network with guided aggregation for real-time semantic segmentation. *Int. J. Comput. Vis.* 129 (11), 3051–3068, <https://doi.org/10.1007/s11263-021-01515-2>.
- Yu, H.-J., Son, C.-H., Lee, D., 2019. Apple leaf disease identification through region-of-interest-aware deep convolutional neural network. *arXiv Preprint: 1903.10356*. <https://doi.org/10.48550/arXiv.1903.10356>.
- Zanetti, M., Samoggia, A., Young, J., 2020. Fruit sector strategic management: An exploration of agro-food chain actors' perception of market sustainability of apple innovation. *Sustainability* 12, 16, <http://dx.doi.org/10.3390/su12166542>.
- Zhang, X., Li, D., Liu, X., Sun, T., Lin, X., Ren, Z., 2023b. Research of segmentation recognition of small disease spots on apple leaves based on hybrid loss function and CBAM. *Front. Plant Sci.* 14, 1175027, <http://dx.doi.org/10.3389/fpls.2023.1175027>.
- Zhang, S., Wang, H., Huang, W., You, Z., 2018. Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG. *Optik* 157, 866–872, <http://dx.doi.org/10.1016/j.jjleo.2017.11.190>.
- Zhang, S., You, Z., Wu, X., 2019. Plant disease leaf image segmentation based on superpixel clustering and EM algorithm. *Neural Comput. Appl.* 31 (2), 1225–1232, <http://dx.doi.org/10.1007/s00521-017-3067-8>.
- Zhang, S., Zhang, C., 2023. Modified U-Net for plant diseased leaf image segmentation. *Comput. Electron. Agric.* 204, 107511, <http://dx.doi.org/10.1016/j.compag.2022.107511>.
- Zhang, Y., Zhou, G., Chen, A., He, M., Li, J., Hu, Y., 2023c. A precise apple leaf diseases detection using BCTNet under unconstrained environments. *Comput. Electron. Agric.* 212, 108132, <http://dx.doi.org/10.1016/j.compag.2023.108132>.
- Zhang, W., Zhou, G., Chen, A., Hu, Y., 2022. Deep multi-scale dual-channel convolutional neural network for Internet of things apple disease detection. *Comput. Electron. Agric.* 194, 106749, <http://dx.doi.org/10.1016/j.compag.2022.106749>.
- Zhang, L., Zhou, G., Chen, A., Yu, W., Peng, N., Chen, X., 2023a. Rapid computer vision detection of apple diseases based on AMCFNet. *Multimedia Tools Appl.* 82, 44697–44717, <http://dx.doi.org/10.1007/s11042-023-15548-x>.
- Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid scene parsing network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 6230–6239, <http://dx.doi.org/10.1109/CVPR.2017.660>.
- Zheng, S., Lu, J., Zhao, H., Zhu, X., Luo, Z., Wang, Y., Fu, Y., Feng, J., Xiang, T., Torr, P.S., Zhang, L., 2021. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR*, pp. 6877–6886, <http://dx.doi.org/10.1109/CVPR46437.2021.00681>.
- Zhu, F., 2023. AppleLeaf9-Enhanced edition(AL9EE). <http://dx.doi.org/10.6084/m9.figshare.23606010.v1>.
- Zhu, Z., Jiang, M., Dong, J., Wu, S., Ma, F., 2023b. PD-SegNet: Semantic segmentation of small agricultural targets in complex environments. *IEEE Access* 11, 90214–90226, <http://dx.doi.org/10.1109/ACCESS.2023.3284036>.
- Zhu, S., Ma, W., Lu, J., Ren, B., Wang, C., Wang, J., 2023a. A novel approach for apple leaf disease image segmentation in complex scenes based on two-stage DeepLabv3+ with adaptive loss. *Comput. Electron. Agric.* 204, 107539, <http://dx.doi.org/10.1016/j.compag.2022.107539>.
- Zou, K., Chen, X., Wang, Y., Zhang, C., Zhang, F., 2021. A modified U-Net with a specific data argumentation method for semantic segmentation of weed images in the field. *Comput. Electron. Agric.* 187, 106242, <https://doi.org/10.1016/j.compag.2021.106242>.
- Zou, K., Wang, H., Yuan, T., Zhang, C., 2023. Multi-species weed density assessment based on semantic segmentation neural network. *Precis. Agric.* 24 (2), 458–481, <https://doi.org/10.1007/s11119-022-09953-9>.