Digital plant pathology: a foundation and guide to modern agriculture

Matheus Thomas Kuska1 · René H. J. Heim2 · Ina Geedicke2 · Kaitlin M. Gold3 · Anna Brugger4 · Stefan Paulus2

Received: 24 February 2022 / Accepted: 17 March 2022 / Published online: 28 April 2022
© The Author(s), under exclusive licence to Deutsche Phytomedizinische Gesellschaft 2022

Abstract
Over the last 20 years, researchers in the field of digital plant pathology have chased the goal to implement sensors, machine learning and new technologies into knowledge-based methods for plant phenotyping and plant protection. However, the application of swiftly developing technologies has posed many challenges. Greenhouse and field applications are complex and differ in their study design requirements. Selecting a sensor type (e.g., thermography or hyperspectral imaging), sensor platform (e.g., rovers, unmanned aerial vehicles, or satellites), and the problem-specific spatial and temporal scale adds to the challenge as all pathosystems are unique and differ in their interactions and symptoms, or lack thereof. Adding host–pathogen–environment interactions across time and space increases the complexity even further. Large data sets are necessary to enable a deeper understanding of these interactions. Therefore, modern machine learning methods are developed to realize the fast data analysis of such complex data sets. This reduces not only human effort but also enables an objective data perusal. Especially deep learning approaches show a high potential to identify probable cohesive parameters during plant-pathogen-environment interactions. Unfortunately, the performance and reliability of developed methods are often doubted by the potential user. Gaining their trust is thus needed for real field applications. Linking biological causes to machine learning features and a clear communication, even for non-experts of such results, is a crucial task that will bridge the gap between theory and praxis of a newly developed application. Therefore, we suggest a global connection of experts and data as the basis for defining a common and goal-oriented research roadmap. Such high interconnectivity will likely increase the chances of swift, successful progress in research and practice. A coordination within international excellence clusters will be useful to reduce redundancy of research while supporting the creation and progress of complementary research. With this review, we would like to discuss past research, achievements, as well as recurring and new challenges. Having such a retrospect available, we will attempt to reveal future challenges and provide a possible direction elevating the next decade of research in digital plant pathology.

Keywords Digitalization · Optical sensors · Robots · Imaging · Machine learning · Digital plant pathology

Introduction
The changing attitude of society towards a more sustainable planet, which is nowadays termed as ‘neo-ecology’, is changing our common agriculture in a drastic way. Stockbreeding, crop cultivation and plant protection are critically re-examined in the view of environmental and human protection strategies to meet the standards of the ‘agriculture green development’ (Davies and Shen 2020). Currently, agricultural land covers approximately five billion hectares, which is 38% of the available land on our planet (annual data FAO 2018). Agriculture must be updated in some aspects to meet rigorous environmental protection targets. However, a sustainable increase in productivity is inevitable because human population is growing continuously. Due to...
the COVID-19 pandemic, the proportion of undernourished people even increased from 650 (~8.4%) up to 811 million (~9.9%) (annual data FAO 2021).

To guarantee sufficient food production, unnecessary production loss in agriculture must be avoided. Globally, integrated pest management (IPM) has reduced harvest losses of the five major food crops (i.e., wheat, rice, maize, potato, soybean) to 20–40% which are attributed to plant pathogens and pests (Savary et al. 2019). Unfortunately, most fields are too large for growers to cost-effectively monitor yield-reducing causes, such as diseases, at regular time intervals. In addition, the detection and exact determination are complex. The field of remote sensing offers methods for high temporal- and spatial-resolution monitoring, that can be used to efficiently deploy ground analysis and remediation action to diseased plants before financial losses incur and disease epidemics emerge. The application of remote sensing methods in plant pathology detection is based on the fact that plant pathogens and pests change the way light interacts with leaves and canopies.

Remote sensing, at its core, is the use of non-contact, often optical sensors such as RGB, multi- and hyperspectral, thermal, chlorophyll fluorescence, and 3D-imaging, to obtain information about processes occurring in the natural and artificial landscape. Optical sensors offer the opportunity for non-destructive disease monitoring at different scales (Mahlein 2016). Next to common techniques for plant disease/pest monitoring, which range from molecular assays to smartphone applications, sensors optimize and reduce the human effort of disease detection in the field (Silva et al. 2021). Though, seemingly straightforward, disease detection, using remote sensing methods in the field, can be complex. Plant diseases themselves are complex as well. They often exhibit a heterogeneous distribution within crop stands and are highly dynamic in time and space due to dynamic interactions between living organisms within an ever-changing environment. Some of the most current challenges, research topics and achievements of digital plant pathology are summarized in Fig. 1, from a phytopathology perspective. It should highlight, that the main goal of digital plant pathology must be to manage farmer’s needs.

Therefore, we are aiming at providing a potential new direction for digital plant pathology research. We are taking a look at some milestones of digital plant pathology and explore the state-of-the-art imaging techniques and analysis methods. With this insight, we are creating a snapshot of the current technical state of applied digital plant pathology and ask the question if we already reached the goal of optimizing manual disease detection.

Digital plant pathology

Almost a century ago, in 1927, Neblette showed that aerial photography (RGB) enables disease survey in agricultural crops. In 1933, Bawden discovered in the lab that a black-and-white representation of an infrared photography resulted in high contrasts between necrotic leaf spots caused by potato viruses. The infrared images were compared to panchromatic (i.e., black-and-white images sensitive to all wavelengths of visible light) images and no obvious contrast was visible. When the same was done with tobacco leaves, the opposite happened, and panchromatic images showed the greatest contrast compared to infrared filter images. The differences were explained by the different makeup of the necrotic areas. Necrotic cells in potato contained chemical break-down products while necrotic cells in tobacco were merely dead empty cells that differed in color compared to the rest of the leaf cells.
These findings set the stage for the use of different spectral bands to detect differences in plant health.

Technical development of optical sensors increased and Colwell (1956) remotely determined wheat rust and other diseases of grains by using military helicopters and infrared-filter cameras, as well as a spectrometer at oblique and nadir observation angles. Colwell suggested to test different combinations of spectral bands for disease detection. Based on the literature and his investigated photos and spectral reflectance curves, he proposed a new view on the interaction of light with plants and the assessment and interpretation of crop photos for plant diseases. Colwell contributed an important theoretical framework that is still of utmost importance in digital plant pathology.

Since 2000, the idea of “foliar functional traits” has strongly emerged as a unifying concept in terrestrial remote sensing to better understand both natural variabilities in vegetation function and variability in response to stress (DuBois et al. 2018). Many traits shown to strongly correlate with natural and stress-induced variation in plant function (Wright et al. 2004) can be quantified and mapped with imaging spectroscopy (Townsend et al. 2003; Ustin et al. 2004; Asner and Martin 2009; Ustin and Gamon 2010; Heim et al. 2015; Wang et al. 2020). Originating in terrestrial ecology, the use of spectroscopy combined with chemistry and taxonomy has been coined as “spectranomics” (Asner and Martin 2009, 2016; Zhang et al. 2020). The foundational components of this approach are: (i) plants have chemical and structural fingerprints that become increasingly unique when additional constituents are incorporated (Ustin et al. 2004) and (ii) spectroscopic signatures determine a portfolio of chemicals found in plants (Jacquemoud et al. 1995).

When applied to plant disease, spectranomics allows for accurate and non-destructive detection of direct and indirect changes to plant physiology, morphology, and biochemistry which induces the disease, both pre- and post-symptomatically (Arens et al. 2016; Couture et al. 2018; Fallon et al. 2020; Gold et al. 2020). Beneficial (Sousa et al. 2021) and parasitic (Zarco-Tejada et al. 2018) plant–microbe interactions impact a variety of plant traits that can be remotely sensed. Changes to narrowband wavelengths, have proven valuable for plant disease sensing due to their sensitivity to a range of foliar properties (Curran 1989). The ultraviolet range (UV; 100–380 nm) is influenced by secondary plant metabolites, while the visible range (VIS; 400–700 nm) is influenced by primary metabolites such as pigments. Internal scattering processes and the structure of a leaf alter the near-infrared range (NIR; 700–1000 nm) while chemicals and water show alterations within the short-wave infrared (SWIR; 1000–2500 nm) (Carter and Knapp 2001). This means, that the nutrient content (Gillon et al. 1999; Zhai et al. 2013; Singh et al. 2015; Wang et al. 2016, 2020), water status (Gao 1996), photosynthetic capacity (Oren et al. 1986), physiology (Serbin et al. 2019), phenolics (Kokaly and Skidmore 2015), secondary metabolites (Couture et al. 2013, 2016) and leaf and cell structure (Mahlein et al. 2012; Leucker et al. 2016; Kuska et al. 2015, 2017), which are changed by diseases are displayed in changes of the spectral reflectance. The foundational spectranomics approach offers an explanation as to why sensing technologies are capable of disease detection in the first place. Remote imaging spectroscopy assesses the sum impact of the fundamental biochemical, structural and physiological processes that underlie the diseased plant phenotype (Mahlein et al. 2012; Leucker et al. 2016, Kuska et al. 2017, 2018a, 2018b, 2019, Zarco-Tejada et al. 2018, 2021; Asner et al. 2018; Sapes et al. 2021). Further ranges of the electromagnetic spectrum can also provide interesting information, but often it is not possible to characterize the determined changes to a specific cause (Mahlein 2016; Simko et al. 2016). As an example, infrared (8–12 µm) light can be determined with thermal cameras, which return a “calibrated” temperature of the plant. The temperature of plants correlates very strongly with the transpiration rate. In addition to recording the water balance of the plant or the crop, this enables the detection of potential drought stress before it becomes visible. Although the sensitivity of thermography and chlorophyll fluorescence sensors is very high, both techniques lack of the possibility to differentiate between abiotic or biotic stress and with it of a causal connection to a specific disease (Mahlein 2016; Simko et al. 2016). However, a combination of sensors can indeed enable a specific characterization of plant diseases.

Within the last couple of years, Zarco-Tejada et al. (2018) were able to use a combination radiative transfer and machine learning approach (Hernández-Clemente et al. 2019) to pre-symptomatically detect Xylella fastidiosa infection in olive trees. This was achieved through a combination of hyperspectral NIR, thermal, and solar-induced fluorescence measurements. The authors found that spectral-plant trait alterations in response to X. fastidiosa infection in both spectral stress indicators and pigment degradation traits, particularly the chlorophyll degradation phaeophytinization-based spectral trait (NPQI), were essential for distinguishing asymptomatically infected plants from both symptomatic and healthy plants. Following up on this work, the authors found that NPQI was only indicative of asymptomatic X. fastidiosa infection in irrigated almond groves. This eventually led to the discovery of the existence of divergent pathogen-and host-specific spectral pathways in response to abiotic and biotic stresses that yield a similar visual manifestation (Zarco-Tejada et al. 2021). Even though both drought and bacterial infection cause the plant to wilt, the mechanisms by which they do so are different, and this difference could be captured with spectroscopy. The authors then used the thermal crop water stress index (CWSI) to uncouple the confounding interaction to improve their misclassification.
accuracy from 37% and 17% to 6.6% and 6.5%, respectively. By assessing spectral trait measurements that captured the underlying physiochemical origin of their diseased plant phenotype, the authors were able to develop a robust disease detection and differentiation methodology for mapping asymptomatic X. fastidiosa infection in multiple crops at scale. This success bolsters and lends hope to ongoing investigations that seek to detect diseases in real-world, multi-stress environments (Fig. 1).

Disease management in the field: can spectral imaging provide the required digital information to control plant diseases and pests?

For disease management, weather-based consultation and forecasting systems (e.g., proPlant Expert.com; RANTISMA), enable the best plant protection measures by their warning services of appearing pests and diseases, since the early 1990s (Newe et al. 2003). The manual field check by the farmer is still necessary, but with digital consulting systems, time management and the process for a successful plant protection measures is optimized (summarized in Damos 2015). However, many techniques and methods are still labor-intensive, and therefore, further progress is necessary. Nilsson (1995) already concluded in his review that remote sensing offers a wider range of sensors and application scales ranging from satellites to ground-based platforms. Nevertheless, depending on the scale, pre-symptomatic and disease-specific detection, as well as the influence of the environment remained a major challenge (Mahlein et al. 2012). This is based on the fact, that plant–microbe interactions are subtle changes in biochemistry and structure. The interactions can be described in compatible (plant pathogenesis) and incompatible (plant resistance response) interactions. To differentiate pathogen attack symptoms, resistance reactions, abiotic stress and spectral signatures of healthy leaves, each of these states had to be characterized in detail (Carter and Knapp 2001). Multi- and hyperspectral imaging is the preferable technique to study such interactions from the cell level to the canopy (Bohnenkamp et al. 2019a, b, 2021).

Variances within and between spectral reflectance signatures were already remotely determined with Landsat-2 imagery. It was used to monitor an epidemic in Pakistan in the late 1970s, the first-ever use of space-borne sensing to monitor disease (Nagarajan et al. 1984). However, a better spatial resolution was needed to precisely explore infections in the field, especially to characterize a pathogen. As the equipment to provide higher resolution was still in an early development stage, spatial resolution, spectral resolution, and costs were closely related. For instance, the amount of generated film could not be stored at reasonable costs, was tedious to analyze as human raters had to screen the images, and no computers were available to perform pixel-wise calculations. The overall progress of remote sensing for abiotic and biotic plant stress was summarized by Jackson (1986).

Two decades later, Chaerle et al. (2007) analyzed resistant tobacco plants, and those susceptible to the tobacco mosaic virus; also, they looked at Cercospora beticola on sugar beet using thermal imaging and chlorophyll fluorescence. They enabled a pre-symptomatic detection and indicated that their studied plant-pathogen interactions could be distinguished. Next, studies using hyperspectral imaging showed that it was possible to discriminate and characterize symptoms of sugar beet diseases, such as Heterodera schachtii, Rhizoctonia solani, Cercospora beticola, Uromyces betae and Erysiphe betae (Hillnhütter and Mahlein 2008; Mahlein et al. 2010; Hillnhütter et al. 2012). Hyperspectral imaging (HSI) further enabled the research community to get a deeper understanding of plant-pathogen interactions. Leucker et al. (2016) were able to display different disease severities of sugar beet leaves inoculated with C. beticola caused by quantitative trait loci (QTL). The increase in phenolic compounds and structural discontinuities caused by tissue collapse, in response to fungal toxins, explain the substantial decrease in reflectance of QTL leaves. Using HSI in the SWIR-range, a variety of micro- and macronutrients such as nitrogen, magnesium, sodium, iron, or copper could also be identified in corn and soybeans undergoing water stress (Pandey et al. 2017). In addition, HSI can be extended to all parts of the plant. For example, Alisaac et al. (2019) showed that HSI of wheat spikelets infected by Fusarium head blight allows for the identification of mycotoxins which was confirmed by quantification of fungal DNA.

Importantly, HSI comes with numerous advantages compared to classical visual monitoring or other analytical methods. It can be applied at different scales—from the cellular level for investigating plant tissue in combination with microscopes, over the individual plant scale in greenhouses or climate chambers, to the canopy scale in field applications with cameras mounted on unmanned aerial vehicles or airplanes (Bohnenkamp et al. 2019a, b, 2021; Heim et al. 2019a). However, in all cases, the analysis of HSI data must be done with care as a great complexity results from a triangular relationship between sensor, pathogen, and environment (Fig. 1). This relationship is further complicated by large amounts of often co-linear data (Thomas et al. 2018a). The effective analysis and interpretation of hyperspectral data are limiting factors for an implementation into plant phenotyping or precision agriculture Mahlein et al. (2018). Automated analysis pipelines and sophisticated data mining and machine learning approaches are necessary to “uncover the spectral language of plants”, as it was shown by Wahabzada et al., (2016).
**Data handling and machine learning**

Once imaging has been completed, a data analysis pipeline must be developed and implemented to ensure retrieval of meaningful information. In Fig. 2, we are presenting a workflow diagram that proposes a potential new multidisciplinary workflow for digital plant pathology research. Several requirements are prerequisites and different subsequent or parallel steps are necessary. After successfully measuring plant data, preprocessing needs to be performed. Steps like de-noising, smoothing, calibration, image segmentation, and outlier removing must be added to transfer the image data to features that can be used as input for machine learning routines (Paulus and Mahlein 2020, Behmann et al. 2015). Literature shows not only the importance of this step but also the huge effort that is required for different sensors in greenhouses and in the field (Bohnenkamp et al. 2021, 2019a, b; Thomas et al. 2018b). Hyperspectral measurements either use the raw reflectance signal as input or, for reduction of data complexity, vegetation indices like the NDVI or OSAVI (Bohnenkamp et al. 2019a, b; Rouse et al. 1974, Rondon 1996). Even though data complexity is reduced, vegetation indices retain their predictive power and can be used for phenotyping approaches with comparably low data input and subtle features. An example was shown for light leaf spot on oilseed rape plants (Veys et al. 2019).

Multispectral approaches in the field can be used in similar ways but usually require a much higher effort for registration and data calibration due to the large area of interest and the fact that the environmental conditions are changing during capturing (Tmušić et al. 2020). For such circumstances, it is discussed how 3D imaging can provide necessary information for data calibration (Paulus 2020; Paulus et al. 2014).

Machine learning provides approaches to give meaning to the data. Supervised learning is used to train a classifier to separate different classes of infection or diseases (Rumpf et al. 2010). Therefore, a labeled dataset to train the model is essential. Commonly, the labeled dataset is split into three different subsets including a training set, a validation set, and a test set. The training set is used to generate a model, the validation set is used to validate it and to perform a fine tuning, and the final test set is then used to calculate various accuracy and error metrics. Comparable to conventional data analysis methods where rules are postulated to analyze the data, machine learning enables to learn these rules by the above-mentioned training process. Although data labelling might require intense manual work, these methods enjoy great popularity in plant science.

A type of machine learning algorithm, the neural network, has been rediscovered during the last decade. Invented during the early 1940s (McCulloch and Pitts 1943), this machine learning approach became only popular later with the development of high computational power. Neural

---

**Workflow in Digital Plant Pathology**

- **Experiment Design**
- **Data Acquisition**
- **Data Cleaning**
- **GIS Processing**
- **Differentiate Treatments**
- **Machine Learning**
- **Relevant Signals?**
- **Feature Selection**

**Multidisciplinary teams increase research quality.**

Fig. 2 The workflow for the interpretation of sensor data using machine learning and linking it to biological processes, using supervised learning and feature importance methods, is shown. Adding the biological knowledge to the interpretation of features would allow for a more mechanistic and transparent machine learning approach as is currently the case. Each step in the process is often performed by a single expert. Thus, detailed knowledge of methods—especially in machine learning—is often not available. An approach involving experts from multiple disciplines would improve current workflows.
networks use the captured images for segmentation, classification, or regression tasks (Barbedo 2021). It is to consider that the data, which is used to train these algorithms, must be of high quality and in a proper quantity to realize results with a high accuracy and low aberration.

A deeper insight into the importance of the input variables is also enabled by further supervised machine learning. Adapted algorithms like Boruta or Recursive Feature Elimination (Chen et al. 2020) provide an importance rating for the machine learning features. When used on hyperspectral plant disease data, these techniques can reveal spectral regions of important wavelength for identifying infected plants (Brugger et al. 2021). In contrast to supervised machine learning methods, unsupervised methods do not need any labeled data or data splitting. These clustering approaches like k-means or hierarchical clustering combine data of similar features and thus give semantic to the data by finding patterns of similarity (Wahabzada et al. 2015). However, results are hard to interpret and need then labelled data to be evaluated. Yet, these routines can be used to find groups of similarities, which have not been noticed before.

At this point, biological insight is needed to connect the output of the machine learning methods to plant- and infection processes (Fig. 2). Recently, approaches integrated expert knowledge as active learning processes in the analysis pipeline, this resulted in significantly improved quality and interpretability of machine learning outputs (Schramowski et al. 2020). To exploit such sophisticated data-driven approaches for real applications by agricultural experts, the models must be biological interpretable, which are now known as “white-box machine learning algorithms” (Fig. 1). They earned this name by aiming at being more explainable and transparent for users interested in the underlying cause for algorithmic outputs. The opposite would be the previous type of algorithms known as “black-box algorithms”. Latest developments show publicly available software libraries, such as the caret package (Kuhn 2008), Keras or Tensorflow (Géron 2019). Nevertheless, these models are only powerful through the underlying training data and rely on high-quality annotated data.

**Digitalization in agricultural practice: are robots the better farmer?**

Since the last turn of the millenium, researchers gained confidence in deploying unmanned terrestrial and aerial vehicles (Fig. 1). These could be equipped with reflectance-based sensors for disease detection with enhanced spatial resolutions allowing for better discrimination between biotic and abiotic stress. Some systems reached a work rate of 3 ha/h (West 2003). Still, variations in illumination intensity, sun/sensor orientation, and/or background soil reflection were impairing consistent and high-quality data retrieval. Another problem turned out to be soil dust leading to detection errors and physical damage to the crops through the vehicle itself. Nowadays, automatization, mechatronics, sensors, electrical engineering and artificial intelligence have reached a level that enables a high degree of autonomy for mobile platforms such as drones, cars, and robots (see Fig. 1 “achievements”). In agriculture, autonomous robots, equipped with sophisticated sensor systems, are the next digitalization step for precise fertilization, pesticide spot-spraying and automated mechanical weeding. Automated robotic applications might even offer an alternative for overcoming shortages of human workers, especially for labor-intensive tasks such as harvesting vegetables or manual weeding (Lowenberg-Deboer et al. 2020).

Furthermore, the implementation of automated systems re-designed agricultural production by considering spatial heterogeneities of plant pest distribution or input parameters such as nutrients, water, and agrochemicals (Saiz-Rubio and Rovira-Más 2020; Wegener et al. 2019). The development of robotic applications for crop management differs with respect to the crop type and cultivation system. One example is the usage of UAVs in the field which are releasing *Trichogramma brassicae*, a natural enemy against *Ostrinia nubilalis* (European corn borer), as a biological control in corn plants (Zhan et al. 2021). In contrast to the manual application of “Trichogramma bags”, UAVs enable a fast and practical application in open land.

In the greenhouse, higher levels of automation, such as robotic harvesting of e.g., pepper (Arad et al. 2020) or robotic plant protection measures in tomatoes (Rincón et al. 2020; Cantelli et al. 2019), are already implemented. Field crops bring a variation of challenges as they can be randomly distributed (e.g., cereals) or planted in rows (e.g., corn, sugar beet, cauliflower). A more and more frequently used application is the selective removal of weed within and between crop rows (Bakker et al. 2010) using actuators such as mechanical weeding tools, laser, stampers or milling heads.

Prototypes of these weeding robots have raised public awareness during the COVID-19 pandemic when trained workers for manual weeding were not available (Mitari-tonna and Ragot 2020). A fast development can be seen for weeding robots, in particular for row crops. These robots are commercially available and can be equipped to deal with different working concepts. The first concept depends on a highly accurate GPS positioning of the seed pill (Griepentrog et al. 2006). Precise seeding with just a minimal error is the prerequisite for orientation and an automated weeding. The robots use the weeding tools on the complete field, except for the area around the planted seed. The second concept is independent of the seeding step. Using digital cameras and an adapted vision recognition system mostly
depending on neural networks and a huge underlying training dataset, the robot is able to detect the crop rows and to adapt its position, heading and navigation path. Furthermore, the weeding tools can be positioned in between or across rows (Machleb et al. 2020).

Position-based systems need a highly precise GPS signal mostly coupled to a real-time-kinematic approach (RTK) which needs to be booked at local suppliers. Without this system, a proper operation is not possible. The system is based on pre-learned positions of seeds and is not aware of changes within the plant population. It does not detect if seeds do not germinate, were eaten by animals or rolled to another position during the sewing process. The robot will continue weeding around these spots or uses its tools where it assumes weeds regardless of the actual presence of the seed or seedling. Nevertheless, position-based systems are robust and operate independent of pre-learned image datasets. They need the data of the sewing process, which is commonly performed by specialist machines, which can do this step at high speed for many rows at the same time.

Sensor-based systems can operate on different types of fields independent of the GPS position and the field structure. Camera images were analyzed in an adapted image processing pipeline. Here, the systems need to separate between vegetation and soil and in a second step between crop and weed. Therefore, a machine learning model based on a neural net approach is used which needs to be trained beforehand on datasets with the same crop under various environmental conditions (Bawden et al. 2017). The bigger this training dataset, the better the segmentation of vegetation and soil, or crop and weed, respectively (Baretto et al. 2021). By extending this dataset, the robot can be adapted for usage on different crops.

While the position-based approach is hard to extend to further aspects but easy to adapt on different crops, the sensor-based approach can be extended to aspects of the adapted treatment of different weed types. Furthermore, the generation of weed maps to distinguish different types of weeds, maps for plant properties like biomass, etc. can be extracted from the sensor-based algorithm which can in a later step be used for adapted control of bigger machines in the field. These maps are currently mainly performed by UAVs (Stroppiana et al. 2018). This concept could be a new basis for subtle disease detection in the field (Görlich et al. 2021). Furthermore, position-based systems need a highly precise GPS signal and are robust and operate independent of pre-learned image datasets. They need the data of the sewing process, which is commonly performed by specialist machines, which can do this step at high speed for many rows at the same time.

A similar outlook for farmers is given for spaceborne monitoring since the European Space Agency’s (ESA) Copernicus program launched their satellites SENTINEL-2A in June 2015 and SENTINEL-2B in March 2017. Besides environmental monitoring and vegetation observation, they enable the monitoring of crop diseases and pests. The SENTINEL satellite sensors have a sophisticated resolution of up to 10 m per pixel and a spectral range from 442–2200 nm with a resolution of 12 spectral bands. Free data access is possible using different commercial software as well as with no-charge browser solutions like the EO Browser by the ESA (https://www.sentinel-hub.com). For some plant diseases and pests, the image time span is critical and short-term applications in the field cannot be conducted, which is currently the main drawback of the free satellite data available. This is because the image frequency depends on the revisit frequency of each single SENTINEL-2 satellite, which is 10 days and in the combined constellation 5 days. In addition, cloud cover might block the field of interest during the imaging. Nevertheless, for retrospective field assessments, as well as research investigations and plant breeding processes, spectral images from satellites are a real benefit to map landscapes with relevant crop and cultivation parameters, identify vulnerable spots, assess the vegetation period or conduct measures for future precision field management (Silva 2021, Segarra et al. 2020). Future satellite programs such as Landsat NeXt will likely unlock new applications and research directions (National Aeronautics and Space Administration, 2021).

One of these applications could be an extension of projects trying to improve the protection of water bodies from unwanted plant protection chemical run-off. Farmers of Germany and Norway have already access to the H₂Ot-Spotmanager (http://synops.julius-kuehn.de). It calculates the environmental risk for waterbodies and their living organisms based on updated satellite, weather, and chemical data. Such applications show the manifold opportunities of satellite data even with a resolution that cannot represent a single plant. Plenty of commercial field management programs which use satellite data are already available (e.g., Xarvio Field Manager, 365 FarmNet, FarmERP, Farmlogs, Agworld, AgriWebb). In these programs, farmers give access to their field data or their whole field index. These data are combined with weather and satellite data to give the farmer a complete overview and information (e.g., about plant nutrition, water status, plant health status,
and necessary protection) around the growth of their crop. Robots and satellite applications entail a large potential for plant protection as these machines integrate optical sensors for monitoring, highly trained and efficient models for detection and they are able to carry different actuators for adapted applications in the field (Balafoutis et al. 2020). In combination with weather-based forecasting systems and information platforms/databases (e.g., EPPO Global Database, ISIP, Animal and Plant Health Service (USDA)) machines could be trained soon to generate computer-based solutions and consulting before and during the vegetation season.

Outlook

As highlighted in this review, the achievements in digital plant pathology are great, but the potential is even greater. To exploit the full potential, the state-of-the-art must be regularly questioned while new challenges need to be defined and solved (Fig. 1). Because pathosystems can be very specific and complex, existing techniques must be critically evaluated and calibrated according to each pathosystems details. Generalized frameworks and models are necessary, which are intuitive and accessible for the farmer. To develop generalized models, a global database with spectral disease and plant spectra, could be a great foundation. An example of such a database from another field is the TRY plant trait database (www.try-db.org; Kattge et al. 2020). Challenges of such a collection of spectral data could involve having a standardized approach to clean and upload data (Paulus and Mahlein 2020). Access to the database must be simple, contributions should be acknowledged, data storage must be sustainable, and data curation for years or decades should be funded. Also, linking sensor type, ambient conditions and other necessary metadata to the uploaded dataset should be a requirement. Currently, many publications are presenting analysis pipelines on few, isolated databases (e.g., Plant Village Data; https://www.kaggle.com/emmarex/plantdisease) that have no relation to the complex situation experience in the field. Often, algorithms are not new, and the biological interpretation is missing. This, however, should be the prerequisite for novel publications in the realm of digital plant pathology. Interlocking the complex aspects of phytopathology, sensors, and machine learning is needed. A global database could help to capture and disentangle this complexity.

Unfortunately, concepts using optical sensors for plant disease detection in the field are not yet established, or are still in its infancy, for them to be integrated into decision-supporting solutions. While many calls exist for conducting interdisciplinary research to solve the remaining and persistent challenges, guidance in the form of funding or academic positions for this type of research rarely exists (Heim et al. 2019b; Bock et al. 2020; Brown et al. 2015). Therefore, we suggest the following research and action steps support the development and application of digital plant pathology in the field:

- Conferences for the development of an international spectral database (ISD) of the global main crops.
- Obligation to provide (image) data for publication and inclusion into the ISD.
- Investigation of the influence of abiotic factors on collected data.
- A standardized framework for the collection of remote sensing data, including metadata on ambient and sensor conditions, and sufficient ground reference data.
- Investigation into scale independence of spectral information.
- Enable machine communication (sensor, platform, computer, analysis software) on a common software basis.
- Proof of concept for field applications based on ecological and economic standards.

Digital plant pathology, as well as the whole digitalization of agriculture, will change farmers’ identity, skills, and work (Klerkx et al. 2019; Zolkin et al. 2021). To effectively implement digital technologies in practical agriculture, educated and trained farmers, as well as local consultants, with a commitment to new digital technologies, are required. Until this adaptation happens, these technologies will only become available to farmers via companies, start-ups, or the advisory services. Independent of the transfer from research to application, the common goal must be highly precise plant protection measures and higher performativity without affecting the sustainability of the natural environment.

Acknowledgements This study was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC 2070-390732324. The authors thank Anne-Katrin Mahlein and all other colleagues involved in research for helpful discussion and collaboration.

Author contributions MTK and RHJH have contributed equally to this manuscript. MTK, RHJM and SP wrote the manuscript. IG revised the final manuscript. KG and AB helped to develop the topic. All authors read and approved the final manuscript.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Albiero D, Garcia AP, Umezuz CK, de Paulo RL (2022) Swarm robots in mechanized agricultural operations: a review about challenges for research. Comput Electron Agric. https://doi.org/10.1016/j.compag.2021.106608
Brugger A, Schramowski P, Paulus S, Steiner U, Kersting K, Mahlein AK (2021) Spectral signatures in the UV-range can be combined with secondary plant metabolites by deep learning to characterize barley–powdery mildew interaction. Plant Pathol 70:1572–1582. https://doi.org/10.1111/ppa.13411

Cantelli L, Bonaccorso F, Longo D, Melita CD, Schillaci G, Muscato G (2019) A small versatile electrical robot for autonomous spraying in agriculture. AgriEngineering 1(3):391–402. https://doi.org/10.3390/agriengineering1030029

Carter GA, Knapp AK (2001) Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. Am J Bot 88:677–684. https://doi.org/10.2307/2657068

Chaerle L, Lenk S, Hagenbeek D, Buschmann C, Van der Straeten D (2007) Multicolor fluorescence imaging for early detection of the hypersensitive reaction to tobacco mosaic virus. Am J Plant Physiol 164(3):253–262. https://doi.org/10.1016/j.ajpph.2006.01.011

Chen RC, Dewi C, Huang SW, Caraka RE (2020) Selecting critical features for data classification based on machine learning methods. J Big Data. https://doi.org/10.1186/s42483-020-00327-4

Colwell RN (1956) Determining the prevalence of certain cereal crop plants: linking spectral characteristics to stress and chlorophyll content. Nature 132:168–168. https://doi.org/10.1038/132168a0

Damos P (2015) Modular structure of web-based decision support systems for integrated pest management. J Big Data. https://doi.org/10.1186/s40537-020-00327-4

Fallon B, Yang A, Lapadat C, Armour I, Juzwik J, Montgomery RA, Cavender-Bares J (2020) Spectral differentiation of oak wilt from foliar fungal disease and drought is correlated with physiological changes. Tree Physiol 40:377–390. https://doi.org/10.1093/treephys/tpaa005

Food and Agriculture Organization of the United Nations (2018) FAOSTAT—Agriculture. Available online at http://faostat.fao.org/faostat/services?subset=agriculture.

Food and Agriculture Organization of the United Nations (2021) FAOSTAT—The world is at a critical juncture. Available online at https://www.fao.org/state-of-food-security-nutrition.

Gao B (1996) NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens Environ 58:257–266. https://doi.org/10.1016/S0034-4257(96)00067-3
Newe M, Meier H, Johnen A, Volk T (2003) proPlant expert.com – an online consultation system on crop protection in cereals, rape, potatoes and sugarbeet. Bull OEPP/EPPO Bull 33:443–449. https://doi.org/10.1111/j.1365-2338.2003.00678.x

Nilsson H (1995) Remote sensing and image analysis in plant pathology. Annu Rev Phytopathol 33:489–528. https://doi.org/10.1146/annurev.phytop.33.091995.002421

Oren R, Schulze ED, Matyssek R, Zimmermann R (1986) Estimating photosynthetic rate and specific carbon gain in conifers from specific leaf weight and leaf biomass. Oecologia 70:187–193. https://doi.org/10.1007/BF00379238

Pandey P, Ge Y, Stoeger V, Schnable JC (2017) High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. Front Plant Sci 8:1348. https://doi.org/10.3389/fpls.2017.01348

Paulus S (2020) Measuring crops in 3D: using geometry for plant phenotyping. Plant Methods 15:103. https://doi.org/10.1186/s13007-019-0490-0

Paulus S, Mahlein AK (2020) Technical workflows for hyperspectral plant image assessment and processing on the greenhouse and laboratory scale. GigaScience 9:90. https://doi.org/10.1093/gigascience/giaa090

Paulus S, Dupuis J, Riedel S, Kuhlmann H (2014) Automated analysis of barley organs using 3D laser scanning: an approach for high throughput phenotyping. Sensors 14(7):12670–12686. https://doi.org/10.3390/s140712670

Rincón VJ, Grella M, Marucco P, Alcatrão LE, Sanchez-Hermosilla J, Balsari P (2020) Spray performance assessment of a remote-controlled vehicle prototype for pesticide application in greenhouse tomato crops. Sci Total Environ 726:138509. https://doi.org/10.1016/j.scitotenv.2020.138509

Rondeaux G, Steven M, Baret F (1996) Optimization of soil-adjusted vegetation indices. Remote Sens Environ 55(2):95–107. https://doi.org/10.1016/0037-0576(95)00186-7

Rouse JW, Haas RH, Schell JA, Deering DW (1974) Monitoring vegetative systems in the great plains with ERTS. Third ERTS-1 Symposium NASA, NASA SP-351, Washington DC, 309–317.

Rumpf T, Mahlein AK, Steiner U, Oerke EC, Pflumer L (2010) Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. Comput Electron Agric 74:91–99. https://doi.org/10.1016/j.compag.2010.06.009

Sáiz-Rubio V, Rojas-Máñez F (2020) From smart farming towards Agriculture 5.0: a review on crop data management. Agronomy 10(5):641. https://doi.org/10.3390/agronomy10050641

Serbin SP, Wu J, Ely KS, Kruger EL, Townsend PA, Meng R, Wolfe BT, Chlus A, Wang Z, Rogers A (2019) From the arctic to the tropics: multimbiome prediction of leaf mass per area using leaf reflectance. New Phytol 224:1557–1568. https://doi.org/10.1111/nph.16123

Silva G, Tomlinson J, Onkokesung N, Sommer S, Mrisho L, Legg J et al (2021) Plant pest surveillance: from satellites to molecules. Emerg Top Life Sci 5:275–287. https://doi.org/10.1042/etls20200300

Simko I, Jimenez-Berni JA, Sirault XRR (2016) Phenomic approaches and tools for phytopathologists. Phytopathology 107(1):6–17. https://doi.org/10.1094/PHYTO-02-16-0082-RVW

Singh A, Serbin SP, McNeil BE, Kingdon CC, Townsend PA (2015) Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. Ecol Appl 25:2180–2197. https://doi.org/10.1890/14-2098.1

Sousa D, Fisher JB, Galvan FR, Pavlick RP, Cordell S, Giambelluca TW et al (2021) Tree canopies reflect mycorrhizal composition. Geophys Res Lett 48: e2021GL092764. https://doi.org/10.1029/2021GL092764

Stroppiana D, Villa P, Sona G, Ronchetti G, Candigiani G, Pepe M, Busetto L, Migliazzi M, Boschetti M (2018) Early season weed mapping in rice crops using multi-spectral UAV data. Int J Remote Sens 39:5432–5452. https://doi.org/10.1080/01431161.2018.1411569

Thomas S, Kuska MT, Bohnenkamp D, Brugger A, Alisaac E, Wahabzada M, Behmann J, Mahlein AK (2018a) Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. J Plant Dis Prot 125:5–20. https://doi.org/10.1007/s12061-017-0246-4

Thomas S, Behmann J, Steier A, Kraska T, Muller O, Rascher U, Mahlein AK (2018b) Quantitative assessment of disease severity and rating of barley cultivars based on hyperspectral imaging in a non-invasive, automated phenotyping platform. Plant Methods 14:45. https://doi.org/10.1186/s13007-018-0313-8

Tmunić G, Manfreda S, Aasen H, James MR, Gonçalves G, Bendor E, Brook A, Polinova M, Arranz JJ, Mészáros J, Zhuang R, Johansen K, Malbeteau Y, de Lima IP, Davids C, Herban S, McCabe MF (2020) Current practices in UAS-based environmental monitoring. Remote Sens 12:1001. https://doi.org/10.3390/rs12061001

Townsend PA, Foster JR, Chastain RA, Currie WS (2003) Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS. IEEE Trans Geosci Remote Sens 41:1347–1354. https://doi.org/10.1109/TGRS.2003.813205

Ustin SL, Gamon JA (2010) Remote sensing of plant functional types. New Phytol 186:795–816. https://doi.org/10.1111/j.1469-8137.2010.03284.x

Ustin SL, Roberts DA, Gamon JA, Asner GP, Green RO (2004) Using imaging spectroscopy to study ecosystem processes and properties. Bioscience 54:523–534. https://doi.org/10.1641/0006-3568(2004)054[0523:UISTSE]2.0.CO;2

Veyes C, Chatziavgerino F, AlSuwaidi A, Hibbert J, Hansen M, Bernotas G, Smith M, Yin H, Rolfe S, Grieve B (2019) Multispectral imaging for presymptomatic analysis of light leaf spot in oilseed rape. Plant Methods 15:1–12. https://doi.org/10.1186/s13007-019-0389-9

Wahabzada M, Paulus S, Kersting K, Mahlein AK (2015) Automated interpretation of 3D laser-scanned point clouds for plant organ segmentation. BMC Bioinform 16:248. https://doi.org/10.1186/s12859-015-0665-2

Wahabzada M, Mahlein AK, Bauckhage C, Steiner U, Oerke EC, Kersting K (2016) Plant phenotyping using probabilistic topic models: Uncovering the hyperspectral language of plants. Sci Rep 6:22482. https://doi.org/10.1038/srep22482

Wang Z, Wang T, Darvishzadeh R, Skidmore AK, Jones S, Suarez L, Woodgate W, Heiden U, Heurich M, Hearne J (2016) Vegetation
indices for mapping canopy foliar nitrogen in a mixed temperate forest. Remote Sens 8(6):491. https://doi.org/10.3390/rs8060491
Wang Z, Chlus A, Geygan R, Ye Z, Zheng T, Singh A, Couture JJ, Cavender-Bares J, Kruger EL, Townsend PA (2020) Foliar functional traits from imaging spectroscopy across biomes in eastern North America. New Phytol 228:494–511. https://doi.org/10.1111/nph.16711
Wegener JK, Urso LM, Von Hörsten D, Hegewald H, Minßen TF, Schattenberg J, Gaus CC, De Witte T, Nieberg H, Isermeyer F, Frerichs L, Backhaus GF (2019) Spot farming—an alternative for future plant production. J Cultiv Plants 71:70–89. https://doi.org/10.5073/JFK.2019.04.02
West JS, Bravo C, Oberli R, Lemaire D, Moshou D, McCartney HA (2003) The potential of optical canopy measurement for targeted control of field crop diseases. Annu Rev Phytopathol 41:593–614. https://doi.org/10.1146/annurev.phyto.41.121702.103726
Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F et al (2004) The worldwide leaf economics spectrum. Nature 428:821. https://doi.org/10.1038/nature02403
Zarco-Tejada PJ, Camino C, Beck PSA, Calderon R, Hornero A, Hernández-Clemente R et al (2018) Previsual symptoms of Xylella fastidiosa infection revealed in spectral plant-trait alterations. Nat Plants 4:432–439. https://doi.org/10.1038/s41477-018-0189-7
Zarco-Tejada PJ, Poblete T, Camino C, Gonzalez-Dugo V, Calderon R, Hornero A et al (2021) Divergent abiotic spectral pathways unravel pathogen stress signals across species. Nat Commun 12:6088. https://doi.org/10.1038/s41467-021-26335-3
Zhai Y, Cui L, Zhou X, Gao Y, Fei T, Gao W (2013) Estimation of nitrogen, phosphorus, and potassium contents in the leaves of different plants using laboratory-based visible and near-infrared reflectance spectroscopy: comparison of partial least-square regression and support vector machine regression methods. Int J Remote Sens 34:2502–2518. https://doi.org/10.1080/01431161.2012.746484
Zhan Y, Chen S, Wang G, Fu J, Lan Y (2021) Biological control technology and application based on agricultural unmanned aerial vehicle (UAV) intelligent delivery of insect natural enemies (Trichogramma) carrier. Pest Manag Sci 77:3259–3272. https://doi.org/10.1002/ps.6371
Zhang N, Yang G, Pan Y, Yang X, Chen L, Zhao C (2020) A review of advanced technologies and development for hyperspectral-based plant disease detection in the past three decades. Remote Sens 12:3188. https://doi.org/10.3390/rs12193188
Zolkin AL, Burda AG, Avdeev YM, Fakhertdinova DI (2021) The main areas of application of information and digital technologies in the agro-industrial complex. IOP Conf Ser Earth Environ Sci 677:032092. https://doi.org/10.1088/1755-1315/677/3/032092

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.