Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: a case study of lettuce production

Authors

Alan Bauer[1,2,3,+], Alan.Bauer@earlham.ac.uk, orcid: 0000-0002-7443-1511

Aaron George Bostrom[1,+], aaron.bostrom@earlham.ac.uk, 0000-0002-7300-6038

Joshua Ball[1], Joshua.Ball@earlham.ac.uk, orcid: 0000-0003-4840-3768

Christopher Applegate[1], Christopher.Applegate@earlham.ac.uk

Tao Cheng[4], tcheng@njau.edu.cn, orcid: 0000-0002-4184-0730

Stephen Laycock[3], S.Laycock@uea.ac.uk, orcid: 0000-0002-7724-1369

Sergio Moreno Rojas[5], sergio.morenorojas@ GS-Growers.com

Jacob Kirwan[5,*], Jacob.Kirwan@GS-Growers.com

Ji Zhou[1,2,3,*], Ji.Zhou@njau.edu.cn, Ji.Zhou@earlham.ac.uk, orcid: 0000-0002-5752-5524

[1]Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ, UK

[2]Plant Phenomics Research Center, China-UK Plant Phenomics Research Centre, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China

[3]School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ

[4]National Engineering and Technology Center for Information Agriculture, MARA Key Laboratory for Crop System Analysis and Decision Making, Jiangsu Key Laboratory for Information Agriculture, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China

[5]G's Growers Limited, Ely, Cambridgeshire, CB7 5TZ

[+]Joint first author [*]Corresponding author

27 **Corresponding authors**

28 Ji.Zhou@njau.edu.cn and Ji.Zhou@earlham.ac.uk; Jacob.Kirwan@GS-Fresh.com

29

Abstract

Aerial imagery is regularly used by crop researchers, growers and farmers to monitor crops during the growing season. To extract meaningful phenotypic information from large-scale aerial images collected from the field, high-throughput phenotypic analysis solutions are required, which not only produce high-quality measures of key crop traits, but also support professionals to make prompt and reliable crop management decisions. Here, we report AirSurf, an automated and open-source analytic platform that combines modern computer vision, up-to-date machine learning, and modular software engineering in order to measure yield-related phenotypes from ultra-large aerial imagery. To quantify millions of in-field lettuces acquired by fixed-wing light aircrafts equipped with a normalized difference vegetation index (NDVI) sensors, we customised AirSurf by embedding a deep-learning classifier trained with over 100,000 labelled lettuce signals. The tailored platform, AirSurf-*Lettuce*, is capable of scoring and categorising iceberg lettuces with high accuracy (>98%). Furthermore, novel analysis functions have been developed to map lettuce size distribution across the field, based on which associated global positioning system (GPS) tagged harvest regions have been identified to enable growers and farmers to conduct ultra-scale aerial phenotyping as well as precision agricultural practises to improve actual yield and crop marketability before the harvest.

Keywords

Lettuce; ultra-scale field phenotyping; deep learning; image analysis; precision agriculture; AirSurf

Introduction

As an important source of vitamins, minerals, and trace elements, leaf vegetables play crucial roles in human nutrition[1]. Lettuce (*Lactuca sativa* L.), one of the most popular staple vegetable foods, has a wide range of tastes and textures cultivated for diverse customer needs[2]. Recent research also indicates that lettuce consumption has positive effects on the reduction of cardiovascular disease and chronic conditions due to its rich nutrients such as vitamin A, *Beta-carotene*, folate, and iron content[3]. While lettuce is an important and nutritional crop, fluctuating environments can increase the fragility of its production[4]. For example, the bad weather in Spain in early 2017 has led to retail prices of lettuce products nearly tripled in UK supermarkets[5]. Severe weather not only causes supply shortage, but also affects crop quality. According to previous studies on lettuce growth and development[6,7], young plants at newly planted phase (i.e. from cotyledons unfolded to three true leaves stage) require cool and damp weather after the transplantation from the greenhouse to the field, whereas lettuce leaves can rapidly become bitter and inedible if the growth is accelerated by high ambient temperature at the head maturity phase (i.e. the stage before flowering). Because of the dynamic nature of lettuce production, the actual yield of lettuces in commercial operations is only around 70-80% of the planted quantity[8]. Hence, to ensure the consistency of supply and quality, it is important for growers and farmers to closely monitor their crops during key growth stages, so that prompt and reliable crop management decisions can be made under changeable agricultural conditions[9].

Aerial field phenotyping has become a popular approach for monitoring crops in recent years. Because it can acquire a large number of crop imagery in field experiments using visible, thermal, and multi-spectrum sensors, it has been widely applied to breeding, farming and crop research[10]. To ensure high-quality aerial image acquisition, the flight route and altitude need to be pre-determined together with the selection of appropriate imaging sensors[11]. For example, for physiological traits such as vegetative greenness and canopy structure, a high-definition RGB camera is sufficient; however, many vegetation indices rely on multi- and hyper-spectral imaging sensors to assess important traits such as biomass, stress level, and yield potential[12]. Recently, with the development of image stitching algorithms and orthomosaic generation methods, very detailed crop images can be collected by unmanned aerial

77 vehicles (UAVs) and fixed-wing light aircrafts, which enables high-quality field phenotyping and high-
78 throughput phenotypic analysis[13].

79 To extract meaningful phenotypic information from large-scale image datasets, a variety of computer
80 vision[14], machine learning (ML), and deep learning (DL) approaches[15] have been utilised. In recent
81 years, much attention has been paid to ML/DL techniques, based on which computational algorithms
82 and models were built to accomplish tasks such as vision-based feature selection, image object
83 classification, and pattern prediction[16–18]. With adequate training data, suitable learning algorithms, and
84 well-defined predictive outcomes, the integration of computer vision, ML/DL, and newly emerged
85 analytic solutions (e.g. distributed computing) could lead to a step change for plant phenomics research
86 in the near future[19].

87 In this article, we present a new analytic platform called AirSurf developed for ultra-scale aerial
88 phenotyping and yield-related phenotypic analysis. The software platform is open-source and combines
89 normalized difference vegetation index (NDVI) aerial imagery for data collection, computer vision for
90 image processing, deep learning (i.e. convolutional neural networks, CNNs) for crop counting, and
91 supervised machine learning for quality assessment. AirSurf was customised for commercial lettuce
92 production so that it could be used to analyse millions of lettuces across the field. We call the tailored
93 software platform "AirSurf-Lettuce" (AirSurf-L), which embeds a CNN model trained with over
94 100,000 labelled lettuce signals to measure lettuce heads and their plantation layouts using ultra-large
95 NDVI images. After scoring lettuce, unsupervised ML algorithms were used to classify lettuce heads
96 into three size categories (i.e. small, medium and large) to assess lettuce quality. To connect phenotypic
97 analysis with marketability and crop management decisions, a novel function has been developed in
98 AirSurf-L to associate global positioning system (GPS) coordinates in a given field with lettuce size
99 distribution, based on which efficient harvesting strategies could be formed to increase marketable
100 yield. .
101

Materials and methods

NDVI aerial imaging and experimental fields

NDVI correlates well with leaf area index and biomass[20] and hence was chosen for yield-related field phenotyping. The imaging sensor used is an industrial standard camera as previously described[21]. The aerial imaging was carried out by a 'Sky Arrow' light aircraft, the lightest weight class (Very Light Aircraft, VLA) of any commercial aircraft, which let the pilot to fly with very little fuel, less than an average farm vehicle. Using VLA at 1000 feet (around 305 metres) in the sky, vast areas can be covered at a flight speed of 180-200 km/hour, during which the NDVI sensor can gather ultra-scale crop imagery to cover four or five fields in a single flight.

The ultra-large aerial NDVI imagery was acquired routinely (i.e. four-five times per season) by G's Growers, the second largest vegetable grower in the UK. The flying route and the imaging protocol were designed to facilitate cross-site crop assessment and yield prediction (Fig. 1A). In this study, we used a series of collected ultra-large NDVI images (1.5-2GB per image) at 3cm ground sample distance (GSD) spatial resolution, for iceberg lettuces at H1 and H2 stages (i.e. moderate compact and crushable head), before lettuce leaves were largely overlapped. Experimental fields in the study were all located near Ely, Cambridgeshire UK, ranging from 10 to 20 hectares, with between 800,000 and 1.6 million lettuce heads in a single field. One field (Field A) planted with around 1 million lettuce heads was used to explain the analysis workflow and associated algorithms of AirSurf-L in the following sections (Fig. 1B). A high-level manual yield counting was conducted by G's growers' field specialists during the harvest, which was used to verify and improve the platform. Lettuces in subsections randomly selected from Field A were scored manually by laboratory technicians and then used as training datasets for the deep learning model.

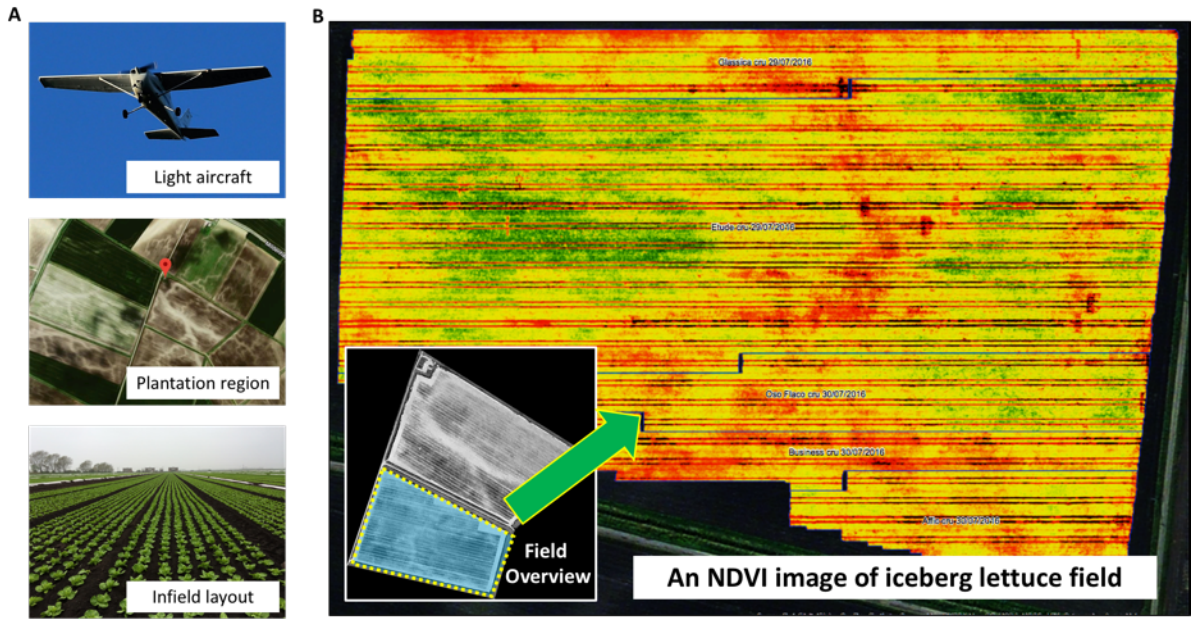


Figure 1: Ultra-scale NDVI aerial imaging accomplished routinely through a fixed-wing light aircraft operated by G's Growers.

(A) The flying route and aerial imaging were designed to facilitate cross-site crop layout assessment and yield prediction. **(B)** A series of ultra-large NDVI images at 3cm GSD spatial resolution were acquired to record 0.8-1.6 million lettuce heads per field, at H1 and H2 stages.

Data construction for training and testing

To generate sound datasets for ML-based image analysis, we randomly selected 60 patches of the field of varying sizes, each containing between 300 and 1,000 lettuce heads. We then manually labelled each lettuce in the selected patches with a red dot (Supplementary Fig. 1). Each labelled lettuce, i.e. a red dot, is identified by a 20x20 pixel bounding box representing a single lettuce head. We then used these bounding boxes, as well as images that did not correspond to a lettuce head, to train a CNN classifier to recognise and separate potential lettuces in the plantation field. The pixels contained within a bounding box were also used for defining lettuce size. A training dataset with over 100,000 20x20 pixel labelled bounding boxes has been created, amongst which 50% are lettuces and the remaining are background signals such as soil, edges of the field, and other non-lettuce objects. Following a standard CNN segmentation approach[22], we designed a non-overlapping sliding window function to go through the whole field to separate foreground and background signals, splitting lettuce and non-lettuce objects.

Training and testing datasets are equally balanced. Validation sets are used alongside training sets to verify the performance of the model, which also prevent overfitting in model training and allow us to fine-tune hyperparameters of different learning layers[23].

The analysis workflow of AirSurf-Lettuce

The analysis of yield-related phenotypes was based on NDVI signals of iceberg lettuces across the field. Figure 2 shows a high-level analysis workflow of AirSurf-L, which consists of five steps: data input, image calibration and pre-processing, ML-based traits analyses, results visualisation, and quantifications of yield-related phenotypes. *Step 1* accepts raw NDVI images as gray-level imagery datasets. As pixels with extremely high NDVI signals usually have overflowed intensity values (i.e. black pixels in Fig. 2A), a pre-processing step (*Step 2*) is designed to calibrate raw NDVI images, so that intensity distribution can be normalised to correct overflowing pixels. At this step, an algorithm called contrast limited adaptive histogram equalization (CLAHE)[24] is applied to increase the contrast between the foreground (i.e. lettuces) and background (e.g. soils) in a given NDVI image (Fig. 2B). Supplementary File S1 provides pseudo code and explanations of the image calibration and pre-processing step to ensure high-quality inputs of the learning model.

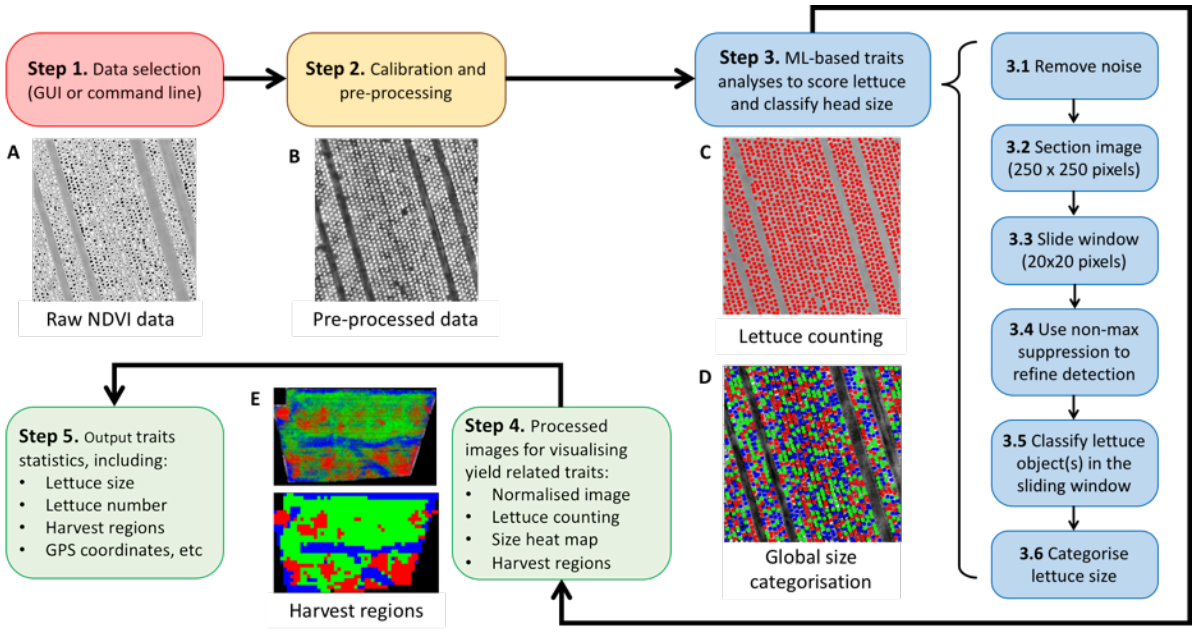


Figure 2: A high-level analysis workflow of AirSurf-Lettuce.

(A) Step 1 accepts raw NDVI images as input imagery data (pixels with extremely high NDVI signals are overflowed). **(B)** Step 2 pre-processes the raw NDVI images to calibrate intensity distribution and correct overflowing pixels. **(C&D)** Step 3 carries out ML-based traits analyses to quantify lettuce number and classify head size in a given NDVI image. (E) Steps 4&5 visualise and export statistics of the traits analyses detection, including yield-related phenotypes such as lettuce counting, size distribution, and harvest regions, and associated GPS coordinates.

Step 3 carries out ML-based traits analyses that quantify lettuce number (Fig. 2C), as well as classify head size (Fig. 2D). It includes six steps: removing noise signals, partitioning a given image into sections (250 x 250 pixels) for local analysis, producing a sliding window (20 x 20 pixels) to traverse within a sectioned image, using non-max suppression to detect lettuces, and classifying recognised lettuces into three size categorises (i.e. small, medium and large). The analysis result is visualised in *Step 4*, where lettuce counting, size distribution map, and GPS-tagged harvest regions are saved as a series of processed images (Fig. 2E). At the final step (*Step 5*), statistics of yield-related traits are exported to a comma-separated values (CSV) file, including lettuce counts per field, lettuce size distribution, lettuce number and size measures within GPS grids, harvest regions, and associated GPS coordinates (Supplementary File S2). To enable users to carry out the above analysis workflow easily, a graphical user interface (GUI) software application has been developed.

AirSurf-Lettuce GUI

The GUI of AirSurf-L (Fig. 3) was developed using the native python GUI package, Tkinter[25], which allows the software application to be executed on different operating systems such as Windows and Mac OS (note: we only provided a packaged .exe executable file on the GitHub). Following the systems design described previously[26], the GUI uses an easy-to-follow approach to implement the phenotypic analysis workflow. The GUI window is divided into two parts: input section and display section. In the input section (dash rectangle coloured red in Fig. 3), a user needs to firstly load an NDVI image, which will be displayed instantaneously in the display section (dash rectangle coloured green), the *original*

tab. Secondly, the user needs to enter GPS coordinates of the field (i.e. the top left corner of the input image, which can be retrieved from the metadata or Google Maps). Thirdly, the user is required to define the rotation value of the input image in degrees in comparison with the north geographical direction, so that GPS calculation can be standardised. Then, the user needs to tell the software whether the input image contains overflown NDVI signals; if so, an extra calibration process will be triggered (Fig. 2, *Step 2*). Finally, after entering a small number of inputs, the user can click the Start button to initiate the analysis workflow.

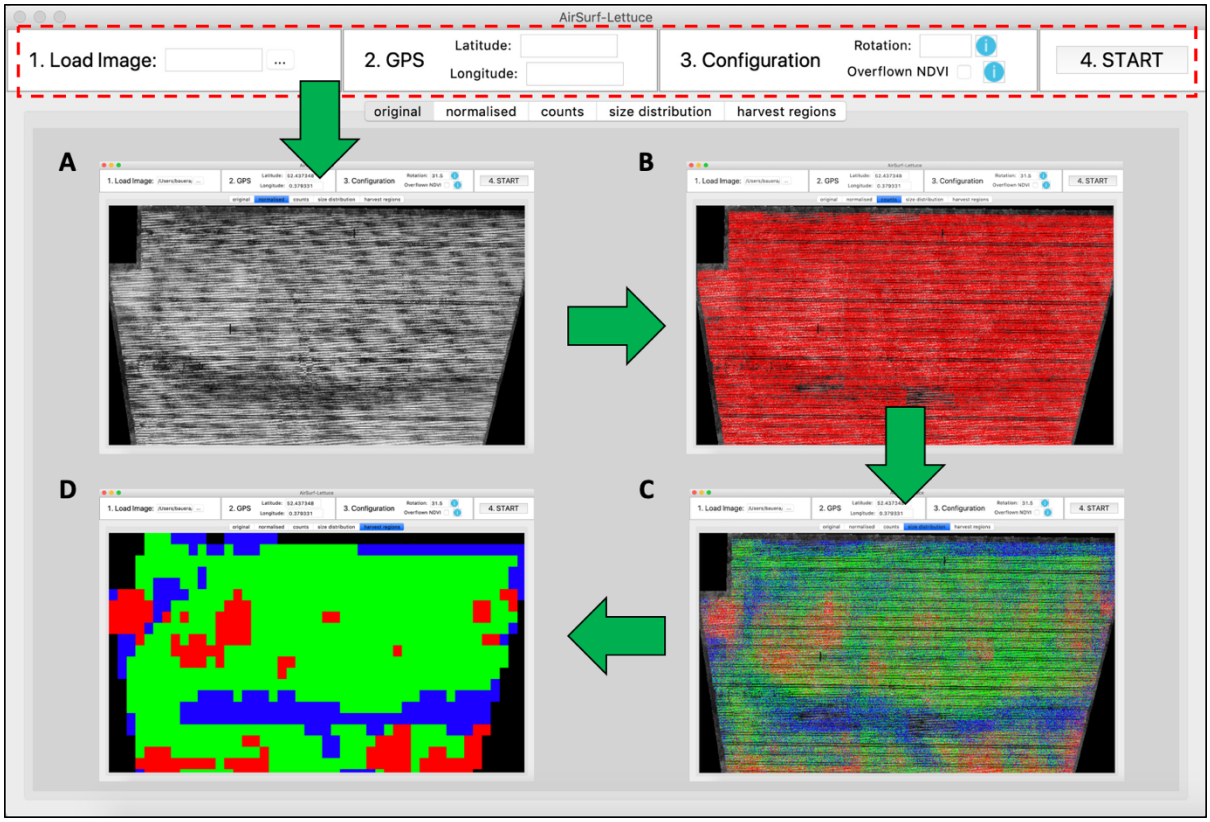


Figure 3: Two sections designed in the AirSurf-Lettuce GUI.

(A) A processed image after pre-processing and calibration (in the normalised tab). **(B)** A processed image after lettuce counting (in the counts tab). **(C)** A processed image after lettuce size categorisation (in the size distribution tab). **(D)** A processed image after identifying harvest regions and commuting GPS coordinates (in the harvest regions tab).

The GUI software follows each step in the workflow to conduct the automated phenotypic analysis. When a step is accomplished, an associated processed image will be displayed in the display section,

showing the result of the intermediate result of the analysis. Four processed images will be presented: a pre-processed and calibrated image (Fig. 3A, in the *normalised* tab), an image after lettuce scoring (Fig. 3B, in the *counts* tab), an image after size categorisation (Fig. 3C, in the *size distribution* tab), and a processed image after identifying harvest regions and their associated GPS coordinates (Fig. 3D, in the *harvest regions* tab). All processed images are saved in a result folder, along with a CSV file that contains trait analysis results (Supplementary File S2).

Neural network architecture

Similar to AlexNet[27], a CNN-based learning architecture was established using the labelled datasets. Figure 4A demonstrates the architecture of the CNN model, including (1) a convolutional (Conv2D) layer with 32 filters and a 3x3 kernel, with a rectified linear unit (ReLU) as the activation function, and batch normalisation to accelerate the learning process to enable higher learning rates[28]; (2) the same block is then repeated together with a max pooling layer to down-sample input using a 2x2 kernel based on the assumption that useful input features could be contained in sub-regions; (3) after that, a second convolutional block is constructed, consisting of a Conv2D layer with 64 filters, a 3x3 kernel, a ReLU activation, and batch normalisation; (4) finally, this block is repeated, followed by another max pooling layer (with a 2x2 kernel) to complete the learning procedure. After the convolutional layers, layers are connected to a fully connected layer of size 512, which is followed by a dropout layer with a 50% chance. To complete the learning architecture, a binary output generates the probability of whether a given bounding box (20x20 pixels) contains a lettuce signal. If the probability equals or is close to 1.0 (i.e. 100%), it indicates that it is highly likely that the bounding box contains a complete lettuce head (Fig. 4B). The above architecture is commonly applied to vision-based object detection problems[29]. The training and validation accuracy and loss curves are reported in Figure 4C, showing that the model converges in only 10 epochs. More importantly, to avoid overfitting, the stopping criterion was designed to guarantee the validation accuracy is higher than the training accuracy, ensuring the generalisation of the learning model. To avoid the overfitting issue of our model, the labelled data was also divided equally into train and validation sets, when training the model.

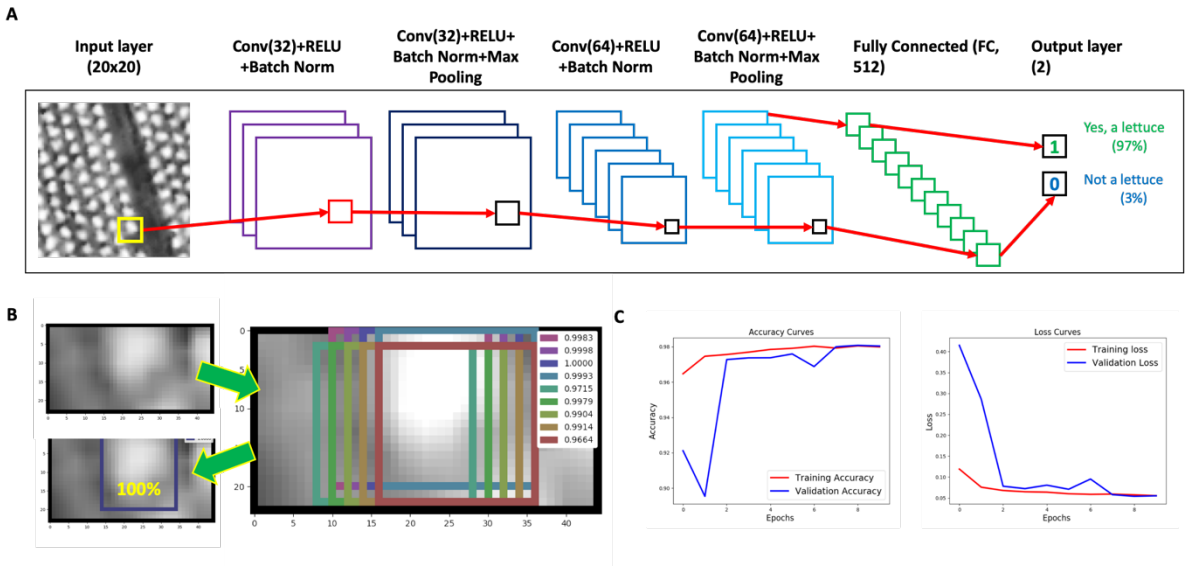


Figure 4: A CNN-based learning architecture established for lettuce counting.

(A) The architecture of the trained CNN model, which generates a binary output representing the probability of whether a yellow bounding box contains a lettuce signal. **(B)** If the probability is close to 1.0, it indicates that it is highly likely that the bounding box encloses a lettuce. **(C)** The training and validation accuracy and loss curves of the model.

The architecture is shallower than AlexNet and other modern deep learning architectures for a number of reasons: (1) the size of our dataset is relatively small for establishing very deep learning networks; (2) our target is a binary classification problem (i.e. whether or not a given bounding box contains a whole lettuce head), opposing to ImageNet classification tasks; (3) larger and deeper neural networks require more time to train, which can be slower to execute and not feasible for prompt decision making requirements in precision agriculture.

Size categorisation algorithm

After AirSurf-L identifies bounding boxes containing lettuce heads, we employed an unsupervised ML approach to categorise lettuce into three sizes: small, medium and large. The algorithm can be easily changed to classify more size categories, if required. Pixels in the bounding box region are extracted and then NDVI values of all the pixels are put into bins. The histogram included 10 bins that spread across the NDVI value range (0-255). We included two important aspects when categorising

lettuce sizes: (1) lower NDVI surrounding values do not determine the lettuce size; (2) higher NDVI values are more important in indicating the size. As such, a geometric pattern of NDVI values for each bin was created, i.e. 64, 128, 160, 192, 208, 224, 232, 240, 244, 248, 250, 252, 253, and 254. With these cut-off values, most of the background pixels were captured in the first two bins, along with increasing weight when the values approach 255.

Having transformed the pixel regions into a series of bins, we were able to compare regions and cluster them into three size groups using k-means clustering with k set to three. Then, clustering results are sorted through calculating the dot product between the weight vector and the cluster count vector (based on bins). These sorted values determine which clustering result corresponds to which size, which are then applied to each lettuce detected in the field. Three colours are used to indicate size categories: blue for small, green for medium, and red for large (Fig. 3C).

Results

Counting lettuces with a CNN classifier

After a CNN classifier was trained and the phenotypic analysis algorithms were implemented in the AirSurf-L, we used the software to recognise and classify lettuce signals in ultra-large NDVI images. Initially, a broad range of sizes and orientations of lettuces with varying intensities were captured; however, the software failed to recognise lettuces in very bright regions and overly count lettuces in very dark regions (Fig. 5A), e.g. around 50,000 lettuces were wrongly detected in a one-million-head field (5% counting error). We found that this problem was caused by the trained CNN classifier, because a lettuce head is extremely tiny in an orthomosaic image (e.g. 11,330x6,600 pixels for a 7-hectare field when GSD is 3cm, which contains over half million lettuces) under extreme lighting conditions. To resolve this issue, we have designed a two-step solution: (1) sectioning the whole image into many 250x250 pixels sub-images, and (2) using a fix-sized bounding-box (20x20 pixels) as a sliding window (with a stepping parameter of 5 pixels to reduce the sliding calculation) to prune the detected lettuce objects in each 250x250 sub-image.

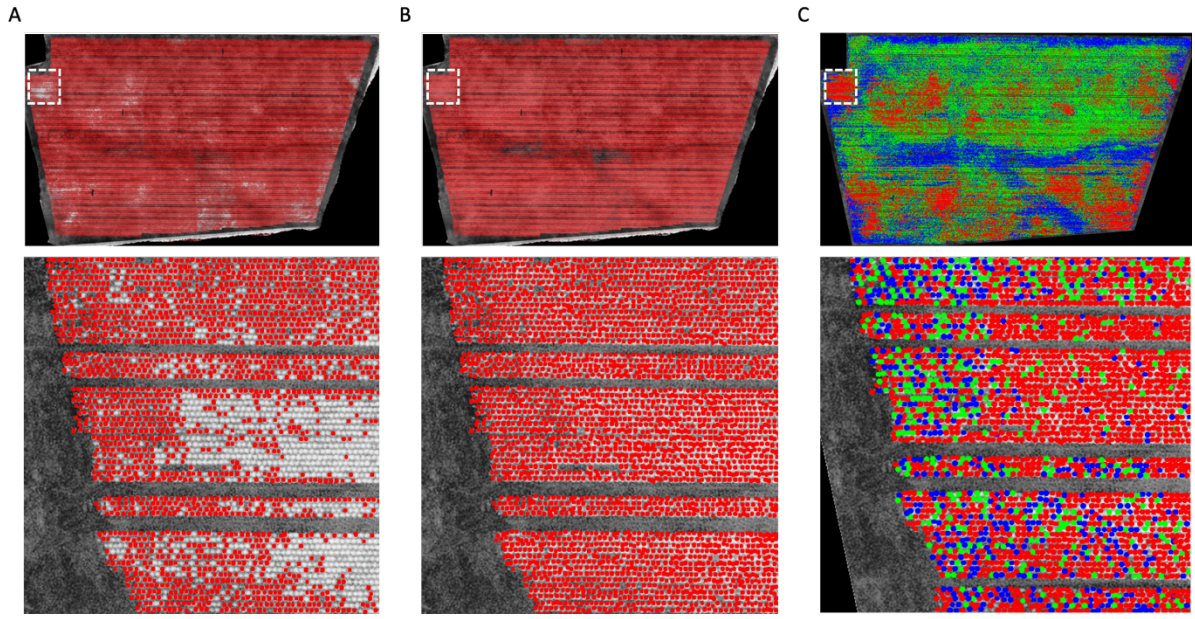


Figure 5: The improved results of the CNN model and the size classification of lettuce heads. **(A)** Wrongly detected lettuces in very bright regions and overly counted lettuces in very dark regions, in a one-million-head field. **(B)** Enhanced training datasets to retrain the model using the online-learning approach, which led to much better detection results. **(C)** A predefined colour code (small is coloured blue, medium is coloured green, and large is coloured red) is assigned to each recognised lettuce head across the field.

Another reason that caused the detection is due to overlapped lettuces as overlapped lettuces could be detected repeatedly by the CNN classifier in a sub-image. Hence, we employed a non-maximum suppression (NMS) algorithm[30] to rectify the detection result. NMS uses probabilities to order the detected lettuce objects. After the 20x20 sliding window is performed and many small patches have been identified, the NMS algorithm computes an overlap coefficient to determine how to retain these patches. As lettuces are relatively well-spaced in the field, patches (i.e. bounding boxes) enclosing a complete lettuce signal are retained, whereas partially covered signals will be removed. To select the best overlap parameter computed by the NMS, a gradient descent method is formulated and explained in Supplementary File S3.

296 *Improved CNN classifier and the size categorisation*

297 Besides the improved computer vision approach, we also enhanced the training datasets by manually
298 labelling an additional 500 lettuce signals within very bright or very dark regions. Then, newly labelled
299 data was inserted into the training datasets to retrain the model through the online-learning approach[31].
300 The improved CNN model (see GitHub repository and Supplementary File S4) was tested on different
301 experimental fields again and has dramatically enhanced the detection result (Fig. 5B).

302 Identified lettuces are individually analysed to determine their associated size category. The size
303 classification is based on intensity and contrast values enclosed by the 20x20 bounding boxes, which is
304 computed using the dot product of the histogram of pixel intensities and a weighted vector towards
305 more pixel-based contrast values. The assumption of this design is that higher NDVI signals likely
306 correlate with higher vegetation indices and hence bigger lettuce heads. The categorisation result of all
307 lettuce heads is clustered into three size groups. Each lettuce is then coloured with a predefined colour
308 code (i.e. small is blue, medium is green, and large is red, see Fig. 5C).

309

310 *A GPS-tagged harvest map*

311 The final phase of the phenotypic analysis is to define harvest regions based on different sizes of
312 lettuces. Using the size distribution map (Fig. 6A), the field is firstly segmented into many small grids
313 based on the optimal GPS resolution determined by the altitude of the aerial imagery (i.e. 3cm GSD, in
314 our case) as well as the size of the harvester machinery used by the grower. After dividing the field into
315 thousands of grids (Fig. 6B), GPS coordinates of each grid are computed and each grid is then coloured
316 with the most representative size category. By combining all coloured grids, a GPS-tagged harvest map
317 is produced, representing harvest regions of the whole field (Fig. 6C). The harvest map can be used for
318 designing harvesting strategies such as guiding a harvester to collect desired sized lettuces or arranging
319 logistics based on the lettuce number and associated size counting. To facilitate precision agricultural
320 practices, a result file (Supplementary File S2) is generated by AirSurf-L at the end of the analysis,
321 containing information of each harvest grid, the associated GPS coordinates, lettuce size and number

counting in each grid. To satisfy different needs for dissimilar requirements, the size of GPS-based harvest grids can be modified manually in the software.

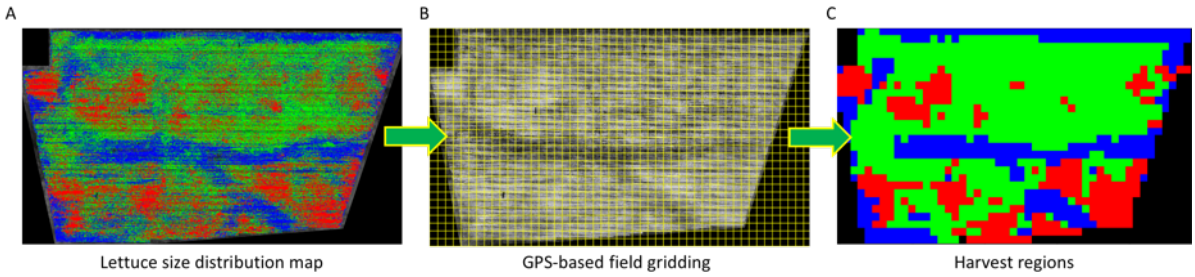


Figure 6: A GPS-based harvest map based on lettuce size classification.

(A) A colour-coded lettuce size distribution map (small is coloured blue, medium is coloured green, and large is coloured red). **(B)** The field is segmented into thousands of grids based on the optimal GPS resolution and the size of the harvester machinery. **(C)** Grids are coloured with the most representative lettuce size category across the image, representing harvest regions of the whole field.

3D visualisation for the harvesting strategy

Figure 7 uses Python-based 3D Matplotlib library[32] to show the GPS-tagged harvest map. When AirSurf-L reads an NDVI image, it computes the number of lettuce heads and associated size categories on the image (Fig. 7A). Then, by combining GPS-based field grids with the representative lettuce size in these grids (Fig. 7B), we produced a Python-based dynamic 3D bar chart script (Supplementary File S5) to present the lettuce number using the z axis, infield layout (both columns and rows) using both x and y axes, and the representative lettuce size using the predefined colours (Fig. 7C). Through the 3D plot, users can zoom into any sub-region of the field to check lettuce number and representative size so that a precise harvesting strategy can be planned accordingly. The overall lettuce number and size counting of the experiment field can also be calculated.

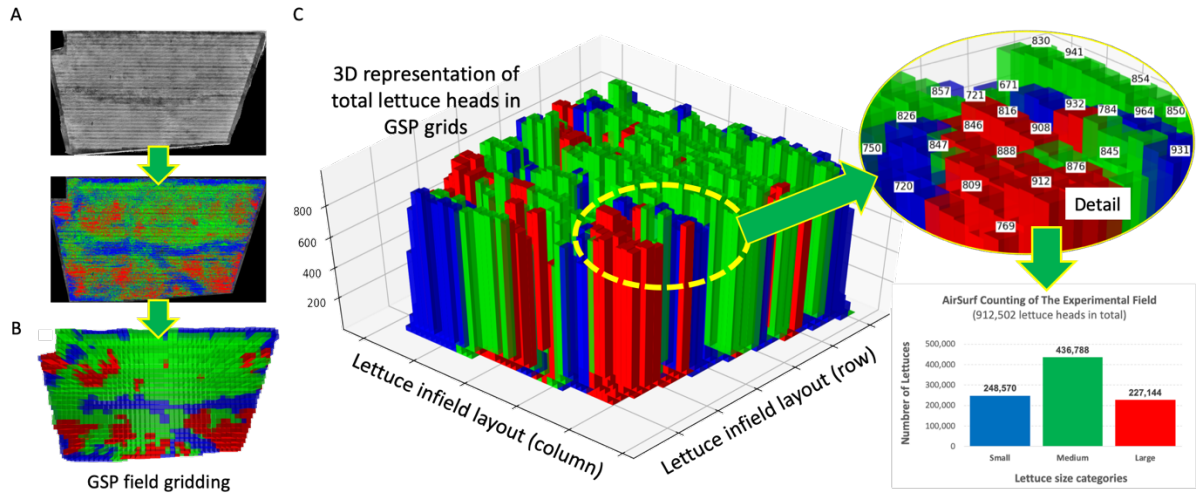


Figure 7: 3D visualisation of lettuce harvest regions.

(A) AirSurf-L reads an NDVI image and exports a lettuce size distribution map, where small lettuce is coloured blue, medium size is coloured green, and large lettuce is coloured red. **(B)** 3D visualising GPS-based field grids to present representative size categories. **(C)** A dynamic 3D bar chart is generated to present the relationship between lettuce number, infield layout, and the representative lettuce size, along with over lettuce number and size quantification.

Validation of AirSurf-Lettuce

To verify AirSurf-L and the soundness the algorithm, we have applied the platform to count and classify lettuce heads in three unseen experimental fields in Cambridgeshire, UK (Figs. 8A-C). These fields contain between 700,000 and 1,500,000 lettuces and are located in different sites around the county. Traits such as the number of lettuces per field quantified by the platform (Fig. 8D) were compared with industrial estimates, showing a low error in lettuce counting (<5% difference). Besides the field-level comparison, we also randomly selected different sizes of subsections in an experiment field to evaluate AirSurf-L. We split these subsections into three sets (i.e. 36 small regions, 21 large regions, and 57 mixed regions), where the small regions have less than 400 lettuces, the large ones contain greater than 900 lettuces heads, and mixed regions contain a variety of lettuce heads. After that, laboratory technicians manually counted lettuce heads within these regions. The correlation between the manual and automated lettuce counting shows that, for the small regions, the difference between the human and automatic counting is approximately 2% (R^2 = 0.978); for the large regions, the average

difference is around 0.8% (R^2 = 0.988); and for mixed regions, the R^2 correlation is over 0.9997. Supplementary Figure 2 and Supplementary File S6 show the correlations between human and automatic counting for all three region groups.

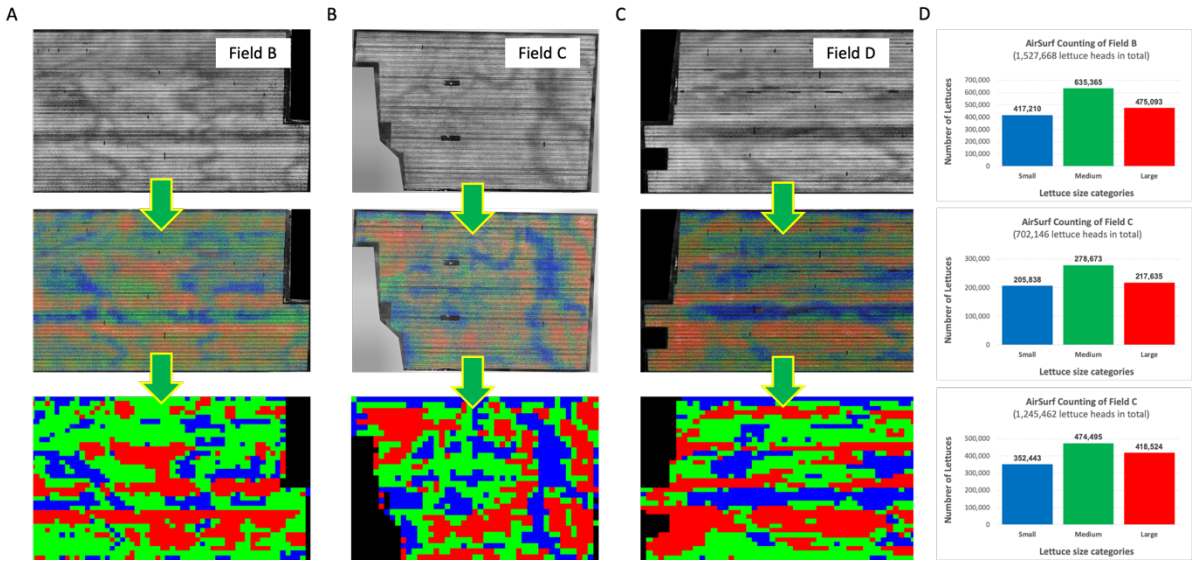


Figure 8: Applying AirSurf-Lettuce to count and classify millions of lettuce heads in three plantation fields across the Cambridgeshire, UK.

(A-C) AirSurf-Lettuce is applied to count and classify millions of lettuce heads (small is coloured blue, medium is coloured green, and large is coloured red), in three plantation fields in the Cambridgeshire, UK. (**D**) The overall quantification of Lettuce heads and size categories in three fields.

Discussion

Traditionally, measuring in-field crops on a large scale is very time-consuming and labour-intensive. It often requires destructive techniques, potentially error-prone manual counting, or estimates of traits that are key to yield production or crop quality[33]. Recent advances in ML/DL and CV techniques have led to an explosion of plant phenomics, which has rapidly improved our abilities in mining phenotypic information from large and complicated phenotyping datasets[34]. New data-driven analytic approaches are also changing plant phenomics – collecting big data (i.e. phenotyping) is no longer the bottleneck, instead how to extract biologically relevant information (i.e. phenotypic analysis) from big data has become the current challenge[35]. Hence, along with the development of aerial imaging and remote sensing technologies, it has become increasingly noticeable that the integration of scalable data

collection, high-throughput phenotypic analysis, and yield predictive modelling are key to crop research and precision agriculture[36].

AirSurf-L introduced here has addressed a specific challenge in ultra-scale aerial phenotyping and precision agricultural management through combining aerial NDVI imagery, CV, ML/DL, software engineering, with commercial lettuce production. The platform automates the measurement of millions of lettuces across the field, which allows us to connect research-based phenotypic analysis with real-world agriculture problems. As a cross-disciplinary project, we have chosen an agile R&D method, because technologies and business requirements were constantly changing in the project. The results generated by AirSurf-L show a strong correlation between automatic counting and specialist scoring ($R^2 = 0.98$). Hence, we are confident that AirSurf-L is capable of assisting fresh vegetable growers and farmers with their phenotyping needs as well as yield-related trait analysis.

Commercial impacts

Commercially, lettuce production offers an attractive economic profitability in comparison to many other Agri-Food businesses[37]. To date, lettuce businesses are worth billions of dollars and employ hundreds of thousands of permanent and seasonal workers globally. European vegetable growers alone produced 2.95 million tonnes of lettuce (and chicory) in 2016, a total annual value of €2.5 billion[38]. Further down the fresh produce supply chain, the planning and efficiency of many essential crop production activities are largely dependent on crop maturity date and the marketability of crops (i.e. the crop quality)[39]. Marketing activities such as logistics, trading, and product marketing need to be organised several weeks before the harvest; moreover, the booking and reservation of crop distribution, agricultural equipment, and associated commercial plans with retails also need to be determined beforehand[40]. By doing so, crop can be harvested at the right time, with maximised yield[41]. Our work contributes directly to lettuce production for improving the actual yield of lettuces and is also capable of reliably quantifying crop quality (e.g. lettuce size), both of which are key to the success of crop production, marketing, and supply chain management.

Machine learning and computer vision in plant phenomics

Another aim of this work is to further ML- and CV-based software solutions in plant phenomics. High-throughput plant phenotyping is a fast-growing research domain, covering many disciplines, from plant breeding, cultivation, remote sensing, to computing sciences[42]. The modulated software development allows us to apply different open-source learning architectures[43] (e.g. the TensorFlow frameworks) and CV algorithms[44] (e.g. Scikit-Image libraries) to AirSurf-L. Notably, it is worth pointing out that we have learned a number of lessons when applying ML/DL and CV in phenotypic analysis: (1) learning algorithms could perform badly if training datasets are not well-labelled and insufficient; (2) although ML/DL algorithms specialised in segmentation and classification if target objects are well-defined, there is still a big gap between object recognition and traits analyses; (3) meaningful phenotypic analysis not only requires sufficient biological understanding to define target traits in a logical manner, but also needs bespoke algorithms to engineer features so that traits can be soundly quantified. Hence, in plant phenomics research, biological questions, analytical solutions, and software implimentation need to be considered collectively in order to address complicated plant phenomics challenges.

Limitations and further development of the platform

Besides the convincing phenotypic analysis results presented in this article, there are still limitations of the platform need to be considered: (1) AirSurf-L has been tested with top-view iceberg lettuces mainly at H1 and H2 stages, which means that analysis error could increase if there are too many overlaps between lettuce heads, e.g. from H3 stage onwards. (2) As AirSurf-L has only been tested with NDVI imagery, it is important to add new functions to the platform to incorporate other vegetation indices measurements through multi- and hyper-spectrum imaging sensors. (3) As precision agriculture management decisions are normally based on imagery, soil and climate conditions, AirSurf's results will be more reliable, if we could include soil conditions for each harvest region and field-level climate conditions. So, results can be compared between sites in multiple years. A potential approach is to incorporate ground-based phenotyping systems such as CropSight[45] to feed environment information to the analysis. (4) The method was tested and validated in lettuce fields in a number of geographic

436 locations following a standard aerial imaging procedure, data collected from different sites via varied
437 aerial imaging strategies (e.g. different angles, altitudes and GSD) could improve the soundness and
438 compatibility of the platform. (5) Key features were constructed by learning algorithms instead of
439 engineered, which make learning models vulnerable when facing up to totally undefined datasets.
440 Hence, ML/DL based phenotypic analysis algorithms need to update with new labelled training data in
441 order to avoid training issues encountered by us. (6) For a field of approximately one million lettuces,
442 it takes about four hours to analyse a field on a decent computer (2.5GHz Intel Core i7, 8 GB memory).
443 Most of the computational time is occupied by the learning model to identify individual lettuces, which
444 can be improved by parallel computing or GPU (graphic processing unit) processing to speed up the
445 phenotypic analysis.
446

447 *Prospects for crop research and precision agriculture*

448 Together with recent advances in multi-scale remote sensing and phenotyping data management[45–47],
449 the platform could be relatively easily expanded to incorporate other crop species such as wheat and
450 rice by retraining the learning models with additional datasets. By doing so, AirSurf could be developed
451 into a more comprehensive analytic platform that will bring great significance to crop production,
452 marketable yield, and precision agriculture management for the Agri-Food sector. For example, the
453 plant density of wheat and rice is closely related to the yield due to its influences on the allocation of
454 water, light and fertilisers, which cannot be quantified using ground-based RGB imagery[48]. Hence,
455 utilising the ultra-scale NDVI aerial imagery and related object recognition methods embedded in
456 AirSurf-L, it is likely to benefit the assessment of sowing performance, emergence rate, and plant
457 distribution. Through a multi-scale field phenotyping approach, breeders and crop researchers could
458 make early predictions of the grain yield of crop genotypes in field experiments.

459 From a precision agriculture perspective, monitoring individual plant such as a lettuce head can enable
460 accurate monitoring of crops during key growth stages across a plantation site. It can provide growers
461 with the real number of crops in the field, based on which yield for harvest availability can be quantified
462 instead of estimated. The calculation of in-field crops can also lead to accurate agricultural inputs,

facilitating automated variable-rate application of fertiliser, weed control, and pesticides through tractor software system with a more precise crop distribution map[49]. Furthermore, the close monitoring of key yield-related traits can be used to guide farmers and growers to reduce variability of agrichemical applications and irrigation in different fields, increasing harvest yield and better operating profit margin[50]. Finally, the AirSurf-L platform utilises existing aerial imagery data routinely performed by growers, which means that no extra data collection cost is required, which is an important factor for new Agri-Tech solutions to be adopted by the Agri-Food sector.

Conclusions

AirSurf-*Lettuce* automatically measures infield iceberg lettuces using ultra-scale NDVI aerial images, with a focus on yield-related traits such as lettuce number, size categories, field size distribution, and GPS-tagged harvest regions. The analysis results are close to the manual counting and can be used to improve existing in-field crop estimates. By monitoring millions of lettuces in the field, we demonstrate the significant value of AirSurf-L in ultra-scale field phenotyping, precise harvest strategies, and crop marketability before the harvest. We believe that our algorithm design, software implementation, the application of ML/DL and CV algorithms, and the cross-disciplinary R&D will be highly valuable for future plant phenomics research that are destined to be more and more challenging. With continuous development work, we are confident that the platform has great potential to support the Agri-Food sector with a smart and precise crop surveillance approach of vegetable crops and therefore lead to better precision agriculture management decisions.

Availability and requirements

Project name: AirSurf-Lettuce with G's Growers

Project home page: https://github.com/Crop-Phenomics-Group/Airsurf-Lettuce

Source code: https://github.com/Crop-Phenomics-Group/AirSurf-Lettuce/

GUI software: https://github.com/Crop-Phenomics-Group/AirSurf-Lettuce/releases

Operating system(s): platform independent

Programming language: Python 3.6

Requirements: Keras, TensorFlow, Skimage, and Numpy.

License: BSD-3-Clause available at https://opensource.org/licenses/BSD-3-Clause

Abbreviations

Comma-separated values (CSV), computer vision (CV), convolutional neural networks (CNNs), deep learning (DL), global positioning system (GPS), ground sample distance (GSD), machine learning (ML), non-maximum suppression algorithm (NMS), normalized difference vegetation index (NDVI), rectified linear units (ReLU), the United Kingdom (UK), and Unmanned Aerial Vehicles (UAVs).

Availability of supporting data

The datasets supporting the results presented here is available at https://github.com/Crop-Phenomics-Group/Airsurf-Lettuce/releases. Source code and other supporting data are also openly available in the GitHub repository.

Author contributions

J.Z., A.G.B., A.B. and J.K. wrote the manuscript, S.M.R. preformed the NDVI imaging. J.K. and S.M.R. provided harvest information and biological expertise. J.Z., A.G.B., and C.A. designed the analysis algorithms. A.B., A.G.B., C.A. and J.Z. implemented the software solution. A.G.B. and A.B. built the deep learning models. J.B. packaged the GUI executables. J.Z., A.B., A.G.B., C.A., S.L. and J.B. tested the software and performed the data analysis. All authors read and approved the final manuscript.

Funding

J.Z. was partially funded by UKRI Biotechnology and Biological Sciences Research Council's (BBSRC) Designing Future Wheat Cross-institute Strategic Programme (BB/P016855/1) to Graham Moore, BBS/E/T/000PR9785 to J.Z.; J.B. were partially supported by the Core Strategic Programme Grant (BB/CSP17270/1) at the Earlham Institute; A.G.B. and C.A. were also partially supported by G's

Growers's industrial fund awarded to J.Z.; A.B. was partially supported by the Newton UK-China Agri-Tech Network+ Grant (GP131JZ1G) awarded to J.Z.

Acknowledgements

The authors would like to thank all members of the Zhou laboratory at EI and Nanjing Agricultural University for fruitful discussions and cross-country collaborations. We thank researchers at John Innes Centre and UEA for constructive suggestions. We gratefully acknowledge the support of NVIDIA Corporation with the award of the Quadro GPU used for this research.

Competing interests

The authors declare no competing financial interests.

References

1 Gubbi J, Buyya R, Marusic S, Palaniswami M. Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions. *Futur Gener Comput Syst* 2013; **29**: 1645–1660.

2 Mou B. Mutations in lettuce improvement. *Int J Plant Genomics* 2011; **2011**: 1–7.

3 Koudela M, Petříková K. Nutrients content and yield in selected cultivars of leaf lettuce (Lactuca sauva L. var crispa). *Hortic Sci* 2008; **35**: 99–106.

4 Tei F, Scaife A, Aikman D. Growth of lettuce, onion, and red beet. 1. Growth analysis, light interception, and radiation use efficiency. *Ann Bot* 1996; **78**: 633–643.

5 BBC. Iceberg lettuces and broccoli rationed as vegetable crisis hits supermarkets. BBC News. 2017; : 1–3.

6 Jie H, Lee Sing Kong. Growth and photosynthetic characteristics of lettuce (Lactuca sativa L.) under fluctuating hot ambient temperatures with the manipulation of cool root-zone temperature. *J Plant Physiol* 1998; **152**: 387–391.

7 Johkan M, Shoji K, Goto F, Hashida S nosuke, Yoshihara T. Blue light-emitting diode light irradiation of seedlings improves seedling quality and growth after transplanting in red leaf lettuce. *HortScience* 2010; **45**: 1809–1814.

8 Ogbodo EN, Okorie PO, Utobo EB. Growth and Yield of Lettuce (Lactuca sativa L .) At Abakaliki Agro-Ecological Zone of Southeastern Nigeria. *World J Agric Sci* 2010; **6**: 141–148.

9 Tudela JA, Hernández N, Pérez-Vicente A, Gil MI. Growing season climates affect quality of fresh-cut lettuce. *Postharvest Biol Technol* 2017; **123**: 60–68.

10 Sankaran S, Khot LR, Espinoza CZ *et al.* Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *Eur J Agron* 2015; **70**: 112–123.

11 Duan T, Zheng B, Guo W *et al.* Comparison of ground cover estimates from experiment plots in cotton, sorghum and sugarcane based on images and ortho-mosaics captured by UAV. *Funct Plant Biol* 2016. doi:10.1071/FP16123.

12 Meyer GE, Neto JC. Verification of color vegetation indices for automated crop imaging applications. *Comput Electron Agric* 2008; **63**: 282–293.

13 Yang G, Yang X, Zhang X *et al.* Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front Plant Sci* 2017; **8**: 1–27.

14 Rousseau D, Dee H, Pridmore T. Imaging Methods for Phenotyping of Plant Traits. In: Kumar S (ed). *Phenomics in Crop Plants: Trends, Options and Limitations*. Springer India, 2015, pp 61–74.

15 Tsaftaris SA, Minervini M, Scharr H. Machine Learning for Plant Phenotyping Needs Image Processing. *Trends Plant Sci* 2016; **21**: 989–991.

16 Burridge J, Jochua CN, Bucksch A, Lynch JP. Legume shovelomics: High-Throughput phenotyping of common bean (Phaseolus vulgaris L.) and cowpea (Vigna unguiculata subsp, unguiculata) root architecture in the field. *F Crop Res* 2016; **192**: 21–32.

17 Bradley D, Xu P, Mohorianu I-I *et al.* Evolution of flower color pattern through selection on regulatory small RNAs. *Science (80-)* 2017; **358**: 925–928.

18 Angermueller C, Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. *Mol Syst Biol* 2016; **12**: 878.

19 Großkinsky DK, Pieruschka R, Svensgaard J *et al.* Phenotyping in the fields: Dissecting the genetics of quantitative traits and digital farming. *New Phytol* 2015; **207**: 950–952.

20 Zaman-Allah M, Vergara O, Araus JL *et al.* Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 2015; **11**: 35.

21 Simms DM, Waine TW, Taylor JC, Juniper GR. The application of time-series MODIS NDVI profiles for the acquisition of crop information across Afghanistan. *Int J Remote Sens* 2014; **35**: 6234–6254.

22 Pound MP, Atkinson JA, Townsend AJ *et al.* Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *Gigascience* 2017; **6**: 1–10.

23 Rosebrock A. *Deep Learning for Computer Vision with Python*. 2017 doi:10.1017/CBO9781107415324.004.

24 Reza AM. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *J VLSI Signal Process Syst Signal Image Video Technol* 2004; **38**: 35–44.

25 Shipman JW. Tkinter 8.5 reference: a GUI for Python. New Mexico, 2013http://www.nmt.edu/tcc/help/pubs/tkinter/.

26 Zhou J, Applegate C, Alonso AD *et al.* Leaf-GP: An open and automated software application for measuring growth phenotypes for arabidopsis and wheat. *Plant Methods* 2017; **13**: 1–31.

27 Alex K, Ilya S, Geoffrey EH. ImageNet Classification with Deep Convolutional Neural Networks. In: *Neural information processing systems*. Neural Information Processing Systems (NIPS): Lake Tahoe, Nevada, 2012, pp 1097–1105.

28 Choromanska A, Henaff M, Mathieu M, Arous G Ben, LeCun Y. The Loss Surface of Multilayer Networks. *arXiv14120233 [cs]* 2014.

29 Matsugu M, Mori K, Mitari Y, Kaneda Y. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks* 2003; **16**: 555–559.

30 Neubeck A, Van Gool L. Efficient non-maximum suppression. *Proc - Int Conf Pattern Recognit* 2006; **3**: 850–855.

31 Hazan E. Introduction to Online Convex Optimization. *Found Trends Optim* 2015; **2**: 151–166.

32 Hunter JD. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 2007; **9**: 90–95.

33 Furbank RT, Tester M, Berry S *et al.* Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci* 2011; **16**: 635–644.

34 Tardieu F, Cabrera-Bosquet L, Pridmore T, Bennett M. Plant Phenomics, From Sensors to Knowledge. *Curr Biol* 2017; **27**: R770–R783.

35 ZHOU J, Tardieu F, Pridmore T *et al.* Plant phenomics: history, present status and challenges. *J Nanjing Agric Univ* 2018; **41**: 580–588.

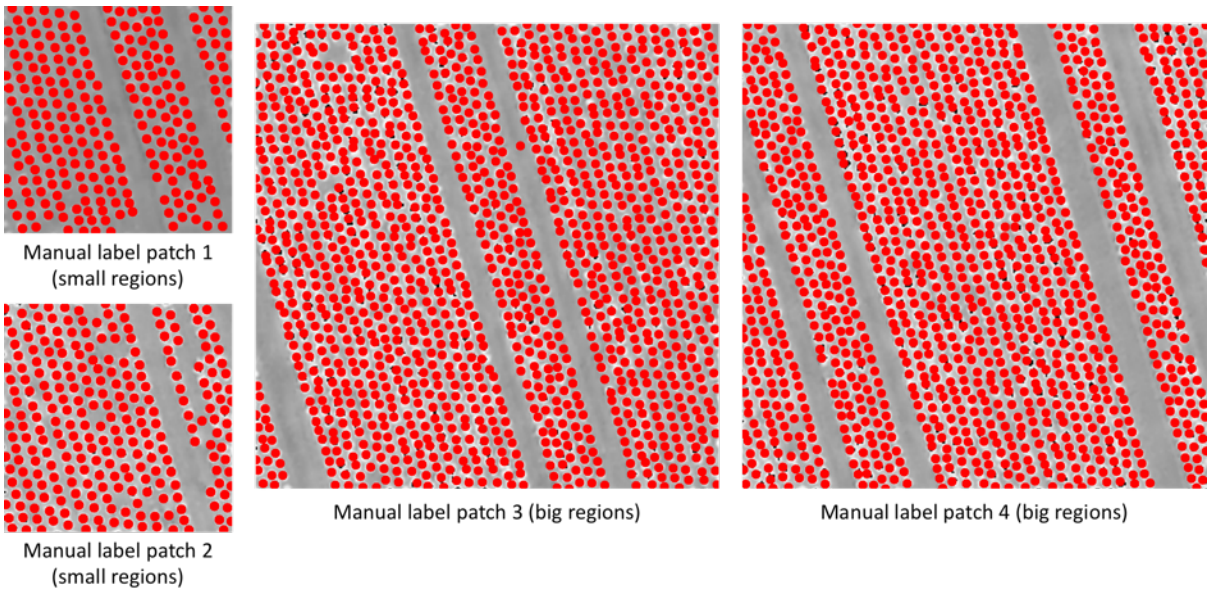
36 Cendrero-Mateo MP, Muller O, Albrecht H *et al.* Field Phenotyping: Challenges and Opportunities. *Terr Ecosyst Res Infrastructures* 2017; : 53–80.

37 Galinato SP, Miles CA. Economic profitability of growing lettuce and tomato in western washington

605 under high tunnel and open-field production systems. *Horttechnology* 2013; **23**: 453–461.
606 38 Mulderij R. OVERVIEW GLOBAL LETTUCE MARKET. Tholen,
607 2016http://www.freshplaza.com/article/163938/OVERVIEW-GLOBAL-LETTUCE-MARKET.
608 39 Shukla M, Jharkharia S. *Agri-fresh produce supply chain management: A state-of-the-art literature*
609 *review*. 2013 doi:10.1108/01443571311295608.
610 40 Lamsal K, Jones PC, Thomas BW. Harvest logistics in agricultural systems with multiple, independent
611 producers and no on-farm storage. *Comput Ind Eng* 2016; **91**: 129–138.
612 41 Schiefer G. New technologies and their impact on the agri-food sector: An economists view. *Comput*
613 *Electron Agric* 2004; **43**: 163–172.
614 42 Singh A, Ganapathysubramanian B, Singh AK, Sarkar S. Machine Learning for High-Throughput Stress
615 Phenotyping in Plants. *Trends Plant Sci* 2016; **21**: 110–124.
616 43 Pedregosa F, Varoquaux G, Gramfort A *et al.* Scikit-learn: Machine Learning in Python. *J Mach Learn*
617 *Res* 2011; **12**: 2825–2830.
618 44 van der Walt S, Schönberger JL, Nunez-Iglesias J *et al.* Scikit-image: image processing in Python.
619 *PeerJ* 2014; **2**: 1–18.
620 45 Reynolds D, Ball J, Bauer A *et al.* CropSight : a scalable and open-source information management
621 system for distributed plant phenotyping and IoT-based crop management. *Gigascience* 2019; **8**: 1–35.
622 46 Reynolds D, Baret F, Welcker C *et al.* What is cost-efficient phenotyping? Optimizing costs for
623 different scenarios. *Plant Sci* 2018; **July**. doi:10.1016/j.plantsci.2018.06.015.
624 47 Neveu P, Tireau A, Hilgert N *et al.* Methods Dealing with multi-source and multi-scale information in
625 plant phenomics : the ontology-driven Phenotyping Hybrid Information System. *New Phytol* 2018; **221**:
626 588–601.
627 48 Liu S, Baret F, Andrieu B, Burger P, Hemmerlé M. Estimation of Wheat Plant Density at Early Stages
628 Using High Resolution Imagery. *Front Plant Sci* 2017; **8**: 1–10.
629 49 Frasconi C, Raffaelli M, Emmi L, Fontanelli M, Martelloni L, Peruzzi A. An automatic machine able to
630 perform variable rate application of flame weeding: Design and assembly. *Chem Eng Trans* 2017; **58**:
631 301–306.
632 50 Brann D, Specialist EG, Sciences SE, Tech V. A Comprehensive Approach Precision Farming :
633 Management is the KEY. *Virginia Coop Ext* 2009.
634

Supplementary Figures

Supplementary Figure 1: Manually labelled lettuces in randomly selected patches using red dots.



Supplementary Figure 2: The correlation between human counting and AirSurf-L scoring.