



HAL
open science

Detecting flowers on imagery with computer vision to improve continental scale grassland biodiversity surveying

Neija Elvekjaer, Laura Martínez-Sanchez, Pierre Bonnet, Alexis Joly, Maria Luisa Paracchini, Marijn van Der Velde

► To cite this version:

Neija Elvekjaer, Laura Martínez-Sanchez, Pierre Bonnet, Alexis Joly, Maria Luisa Paracchini, et al.. Detecting flowers on imagery with computer vision to improve continental scale grassland biodiversity surveying. *Ecological Solutions and Evidence*, 2024, 5, 10.1002/2688-8319.12324 . hal-04593804

HAL Id: hal-04593804

<https://hal.inrae.fr/hal-04593804>

Submitted on 30 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE

Detecting flowers on imagery with computer vision to improve continental scale grassland biodiversity surveying

N. Elvekjaer¹  | L. Martinez-Sanchez¹  | P. Bonnet²  | A. Joly³  |
M. L. Paracchini¹ | M. van der Velde¹ 

¹European Commission, Joint Research Centre (JRC), Ispra, Italy

²AMAP, Université de Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier, France

³INRIA Sophia-Antipolis ZENITH Team, LIRMM, Montpellier, France

Correspondence

N. Elvekjaer

Email: neija.elvekjaer@geo.hu-berlin

M. van der Velde

Email: marijn.van-der-velde@ec.europa.eu

Funding information

JRC Exploratory Research Program through the Rural Refocus Project, Grant/Award Number: 31280; European Union's Horizon 2020 Research and Innovation Program, Grant/Award Number: 863463 and 101017567; IFCA

Handling Editor: Alicia Foxx

Abstract

1. Large-scale biodiversity monitoring is essential for assessing biodiversity trends, yet traditional surveying methods are limited in the spatial/temporal scale they can cover. Recent technological developments have led to computer vision-based species identification tools, such as the Pl@ntNet application. Increasing accuracy of such algorithms presents an opportunity of integrating computer vision into larger monitoring schemes and could lead to automating ground-based evidence provision related to agri-environmental measures (e.g. flower strips, field margins). However, images from surveys or farmer declarations do not live up to the standards of current applications. In order to integrate these automated methods into biodiversity monitoring, more generalized models are needed.
2. We create a dataset using 500 manually delineated images of vegetation patches in European grasslands taken during the Land Use/cover Area Survey (LUCAS) grassland module. We train the Faster R-CNN model to detect and extract individual flower objects. Using this model, we extract the abundance of flowers in an image, analyse their colour distribution, and use the Pl@ntNet application to identify the species of the individual flowers detected.
3. The best model reaches precision and recall of 0.89/0.61 and predicts 1377 flowers on the 100 test images distributed between 10 colours. Using Pl@ntNet, only 52 flowers were identified with a certainty score above 0.5 due to the limitations in image size and quality. Of these flowers, 30% were correctly automatically identified at the species level and 42% at the genus level.
4. The results show that we can automatically extract valuable information on floral abundances, colours, and sizes from images of vegetation patches, though in most cases better images are needed for species identification. Despite limitations with image quality, integrating this workflow into large-scale monitoring could speed up the sampling process and allow for better spatial and temporal data on floral diversity and abundance.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Ecological Solutions and Evidence* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

KEYWORDS

automated species identification, biodiversity monitoring, computer vision, Faster R-CNN, flower, object detection, Pl@ntNet, vegetation survey

1 | INTRODUCTION

Loss of biodiversity and its impact on ecosystem stability and functioning is a major concern worldwide (IPBES, 2019). In farmlands, agricultural intensification has caused significant declines in biodiversity with detrimental effects on biodiversity-mediated ecosystem services such as nutrient cycling, pest regulation, and pollination, resulting in reduced yields (Boetzi et al., 2021; Cole et al., 2020). In the European Union (EU), this decline in biodiversity has triggered several policies and initiatives aimed at preserving farmland biodiversity, such as the EU Pollinators Initiative (European Commission Eu Pollinators Initiative, 2018), the Farm to Fork Strategy target to reduce use and risk of chemical pesticides, and a larger share of the budget available for the environmental dimensions of the Common Agricultural Policy (CAP) for the period 2023–2027. Good agricultural and environmental conditions (GAEC) requirements, ecoschemes, agri-environmental and climate measures can promote the creation and maintenance of biodiversity rich field margins (e.g. flower strips), increase the area of landscape features (e.g. hedges, tree lines), strengthen green corridors within patchy rural landscapes, and manage grasslands in a way that is beneficial for climate (e.g. carbon sequestration) and/or biodiversity. However, monitoring their implementation and maintenance is a challenge and many practises cannot be detected with remote sensing methods (Sima et al., 2020). Furthermore, several studies suggest that the result of the measures is context-dependent and varies between landscapes and regions (Cole et al., 2020; Nilsson et al., 2021; Scheper et al., 2015). Temporal continuity, connectivity, and ecological contrast to the surrounding landscape are important factors for the outcome of such schemes (Boetzi et al., 2021; Nilsson et al., 2021; Scheper et al., 2013, 2015). The effectiveness of measures aiming to increase the availability of floral resources for pollinators, such as those promoting flower strips, supporting environmentally sensitive grasslands, and establishing field margins, can be assessed by not only measuring species abundance but also species composition and presence of key species (Bartual et al., 2019; Cole et al., 2020; Scheper et al., 2015; Sutter et al., 2017). The best implementation might, in fact, be a trade-off between ecosystem services supply and biodiversity conservation (Schaub et al., 2020; Scheper et al., 2013). To get the desired outcome, agri-environmental schemes must be targeted to local conditions, requiring a large amount of knowledge on the ecosystem at a fine temporal and spatial scale. Lack of biodiversity data prevents proper targeting of such management and conservation efforts, reducing potential ecosystem benefits.

Large-scale and accurate biodiversity monitoring is key to understanding ecosystem interactions and the impacts of agriculture (Scholes et al., 2012). However, such monitoring schemes are

labour intensive, may suffer from a lack of temporal/spatial representation as well as surveyor bias. An example of such a survey is the 2018 Land Use/Cover Area frame Survey (LUCAS) grassland module (Sutcliffe et al., 2019), a pilot study aimed at collecting information on the environmental and ecological quality of grasslands and their management throughout the EU. As part of the overall LUCAS campaign in 2018, 2173 points in 26 countries were visited by surveyors and expert botanists. Various information was collected, such as intensity of management, vegetation cover, number of flowering plants, and presence of key species. The final report highlights two problems of this pilot survey: timing and expertise (Oppermann, 2021). A comparison between the data collected by normal surveyors and expert botanists show that the timing of the field visit is very important—for example the species flowering change throughout the growing season. Furthermore, surveyors without expert botanist knowledge had difficulty determining some variables such as vegetation composition and the presence of specific key species. The number of expert botanists is declining, and this taxonomic gap has been recognized as a major obstacle for conservation efforts since the Rio Conference in 1992 (Joly et al., 2019). Future large-scale biodiversity monitoring schemes need to develop new methodologies to sample at an adequate spatial and temporal scale, without relying solely on expert botanists.

1.1 | Computer vision in biodiversity monitoring

Recently, machine learning, especially computer vision algorithms, have been used for species identification of animals (Villon et al., 2020; Weinstein, 2018), insects (Høye et al., 2021), and plants (Joly et al., 2014; Mäder et al., 2021; Mann et al., 2022). These algorithms can mitigate the loss of expert knowledge and the limited spatial/temporal cover of traditional surveying methods (Wäldchen et al., 2018). Although they are far from outperforming the best expert botanists in complex settings (Bonnet et al., 2016), in a fast moving field, they can provide rapid, objective, and scalable species identification, only requiring image input.

These developments go hand in hand with opportunities to collect biodiversity data. Crowd-sourcing and citizen science projects generate and continuously update massive amounts of data by volunteers taking pictures of the biodiversity surrounding them (Boho et al., 2020; Wäldchen et al., 2018). Furthermore, monitoring for compliance assurance, such as in the CAP context (Sima et al., 2020), can require farmers to provide image proof of implemented practises, for example, a picture of a flower strip. For example, farmers in the Thuringia region in Germany can use a plant species identification app, based on Flora Incognita (Mäder et al., 2021), to prove

the occurrences of six plant species out of a list of 44 to receive payments for biodiverse grasslands (Barmeier, 2021). In the future, and given ever-developing sharing mechanisms guaranteeing anonymity, these potentially large amounts of geotagged photos captured through CAP monitoring may contribute to biodiversity monitoring as well.

For computer vision algorithms, it is important to have a precise and representative training dataset. For plant detection and identification, the training data set must capture the huge amount of plant species in the world (400K flowering plants), inter-species variation (some species may look similar), and intra-species variation (the same species may look different depending on growth stage, growth conditions, and disturbances), as well as variations due to image acquisition (image quality, light conditions and viewpoints) (Wäldchen et al., 2018). Current applications for plant species identification use citizen science approaches to generate the extensive data required. However, this approach suffers from high variability in data quality, as it is hard to ensure that the data are collected in a correct and consistent manner. The crowd sourced data is collected by non-experts with varying equipment, expertise and skill and often show significant biases due to (1) the geographical variation in sampling effort (e.g. affected by population density and accessibility) and (2) citizens/non-experts are more likely to miss rare species and sample 'eye catching' species, resulting in long tailed distribution (de Lutio et al., 2021; Jones, 2020) with many observations of common species and few/no observations of rare species. These problems are mitigated to some extent by automatically filtering and relying on data quality checks by a network of experts.

iNaturalist (Unger et al., 2021), Flora Incognita (Mäder et al., 2021), and Pl@ntNet (Joly et al., 2014) are examples of automated image-based species identification applications with global reach that exploit the potential of computer vision and citizen science. In this paper, we use Pl@ntNet, an app developed by a consortium of four French research organizations. Users can send up to 4 images per plant query which returns a list of species classifications and probability scores. Including several images of high quality with focus on different plant organs can enhance the model classification score. The detected species can be verified by a network of experts and, if the images include a location, incorporated into the training dataset as well as the species occurrence data that Pl@ntNet contributes to the GBIF (Global Biodiversity Information Facility). The plant query made by curious citizens becomes part of large species distribution monitoring and useful information for the scientific community (Pl@ntNet Contribution to GBIF, 2024).

Recently, species identification applications have started to include metadata in order to account for the high variability of natural environments. Many applications allow multiple input images with different viewpoints (flowers, leaf, whole plant, etc.) (Rzanny et al., 2019). Others include geo-location of the image, and from the position various environmental conditions such as climate and terrain can be inferred (de Lutio et al., 2021; Terry et al., 2020). Finally, recent studies suggest to take the hierarchical nature of taxonomy into account, possibly using taxonomic knowledge to

infer family-level information of species unknown to the algorithm (de Lutio et al., 2021; Seeland et al., 2019). All this ancillary data is improving the detection algorithm, to the point that it can compete with human experts in simple settings (de Lutio et al., 2021; Jones, 2020; Mahecha et al., 2021; Wäldchen et al., 2018).

This combination of artificial intelligence and increasing amount of image data has been used for species distribution modelling (Botella et al., 2018), extraction of macroecological patterns (Mäder et al., 2021; Mahecha et al., 2021), tracking invasive species (August et al., 2015; Terry et al., 2020), and various conservation projects around the world (Bonnet et al., 2020). Hicks et al. (2021) showed that species recognition can be used to estimate nectar sugar mass, in order to monitor conservation efforts directed towards pollinators. Their computer vision-based approach could cut pollinator-plant survey time per stand of vegetation from hours to minutes. These studies suggest that automated species recognition is mature enough to contribute to large-scale monitoring efforts. Taking advantage of such automated methods and citizen science for data acquisition makes it possible to revisit the sites multiple times during the growing season or expand the spatial scale of the survey. Extended spatial and temporal knowledge on biodiversity patterns will greatly improve the targeting ability of conservation efforts and agri-environmental schemes such as the implementation of green corridors, flower strips, etc.

Current efforts to use computer vision-based methodology in monitoring are limited to identification of specific species using closeup images taken of single flowers for computer vision purposes. However, in field settings, images captured might not be suitable for such identification algorithms due to varying light conditions, picture quality, multiple flowers in the images, etc. In the context of environmental monitoring and assessing the impact of agri-environmental management practices, these algorithms ultimately need to move from the individual flower level towards community- or ecosystem level. In this paper, we aim to develop a generic flower detection model that can bridge this gap. The detection model could automatically generate information on floral resources and diversity at relevant spatial scales currently missing for monitoring floral diversity in farmlands. Using such a model biodiversity information can be extracted from images covering larger vegetation patches/communities and metrics such as flower abundance and colour composition can be obtained in a snapshot. Ideally, the individual flower objects detected by our generic flower model could then be identified using existing identification applications such as Pl@ntNet. Such a framework can scale up if implemented within larger monitoring efforts or within compliance assurance monitoring schemes.

1.2 | Objectives

This paper aims to develop a generic flower detection algorithm and automatically extract information on flower presence, abundance, and diversity, from images taken during the LUCAS grassland survey in 2018. Detailed objectives are:

1. Create a training dataset for object detection of flowers by manually delineated images from the LUCAS grassland survey.
2. Train the Faster R-CNN object detection model to detect flowers on images and evaluate model performance.
3. Extract information from images on flower presence, abundance, and diversity using our flower detection model.
4. Identify the detected flowers using the Pl@ntNet application for species identification.
5. Identify limitations in this methodology and provide recommendations on how this could be adapted to future needs with a view on increasingly automated surveying of flower species and habitats.

2 | MATERIALS AND METHODS

To accomplish our objectives we train a Faster R-CNN model (Ren et al., 2015) to detect flowers in photos. We construct a training dataset from images taken during the LUCAS grassland survey and

label them with the annotation CVAT tool (Sekachev et al., 2020). We detect and extract flowers with the trained model and attempt to identify species by using the Pl@ntNet API (Affouard et al., 2017). Finally, we evaluate the limitations of this methodology and provide recommendations on how to better integrate computer vision-based tools in large-scale biodiversity monitoring of grassland flowering plants. An overview of the workflow can be seen in Figure 1.

2.1 | Image dataset

2.1.1 | LUCAS grassland module

The images used in this study were collected through the LUCAS, coordinated by Eurostat, the statistical office of the EU. The surveys in 2006, 2009, 2012, 2015, 2018, and 2022, cover the extent of the EU with a sampling grid of 2 km². In each campaign, a stratified sample of points is surveyed (e.g. 330k in the 2018

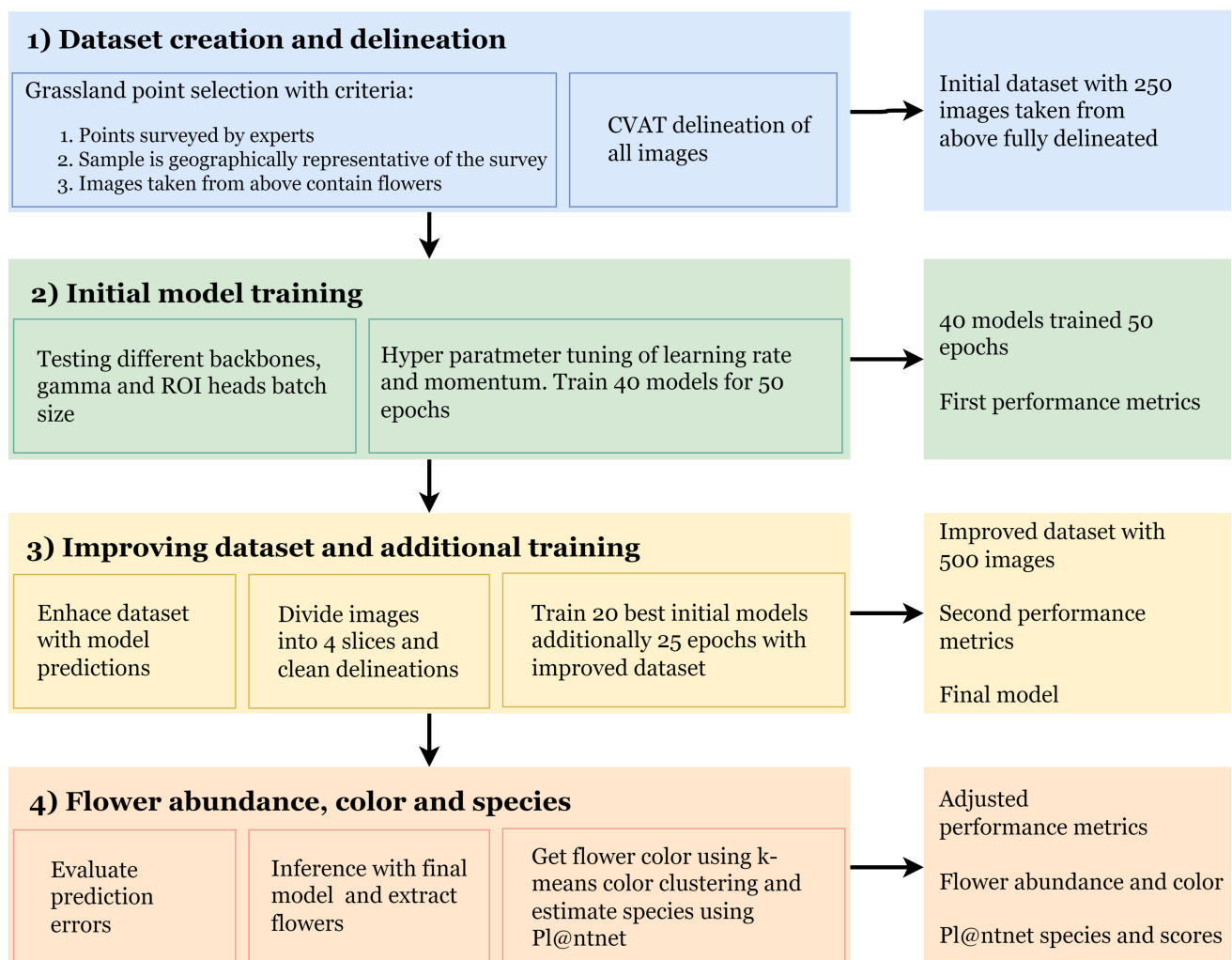


FIGURE 1 Flowchart showing workflow (left) and outcomes (right) of this paper.

campaign); data is collected on land cover, land use, environmental variables, and photos taken in four cardinal directions (d'Andrimont et al., 2020). In 2018, the grassland module was added to the survey, a pilot study to collect detailed information on the environmental and ecological quality of grasslands (Oppermann, 2021). A stratified sub-sample of the LUCAS points was derived covering different grassland regions of Europe. These points were visited in the field within a predefined optimal time frame by both LUCAS surveyors and, in 20% of the cases, expert botanists. The additional visit by experts was done to evaluate the accuracy of the information collected by the surveyors with limited knowledge of grasslands and species identification. 3734 LUCAS points were visited by surveyors and 747 of these were also visited by expert botanists. At each sample point a transect of 2.5×20 m was established and information on environmental and ecological variables was collected, including vegetation structural parameters, presence of key species, and land use practices. All samples include three images, two taken from each end of the transect towards the centre and one top-down image taken of a representative patch of vegetation

at any location within the transect, 1.5 m above the ground (see Figure 2).

2.1.2 | Image selection and processing

A total of 250 points were selected from the LUCAS grassland survey to create an image dataset. A 60/20/20 split was done resulting in 150 points used for training data, 50 points for validation data and 50 points for test data. To ensure high quality of the images taken in the field, only images from the expert surveyed points were used. This was done since we assessed that the experts used better cameras, were better at following the protocol, and used their botanical expertise to capture the flower as best as possible. For a first dataset, we reasoned that these were important parameters. To ensure a good representation of the European grasslands regions defined by the survey, a stratified sub-sample was selected based on the geographic distribution of the surveyed points within these regions. Photos taken from above at these

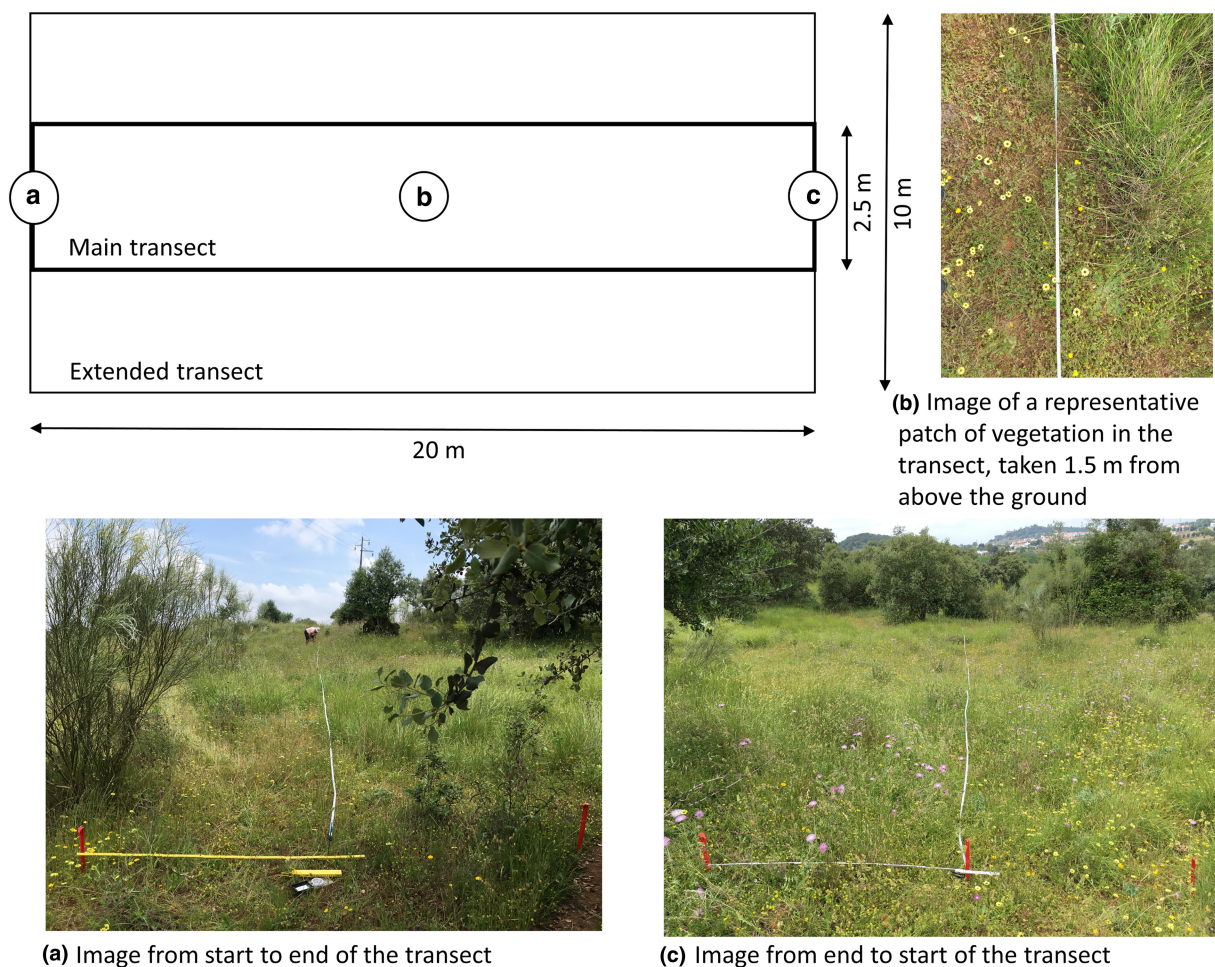


FIGURE 2 Overview of the protocol and pictures taken during the LUCAS grassland module. Illustrating the transect and the three pictures taken during the survey. (a) Image from start to end of the transect. (b) Image of a representative patch of vegetation in the transect, taken 1.5 m from above the ground. (c) Image from end to start of the transect.

selected points were used for the image dataset and the presence of flowers on the images was visually confirmed. Hereby, a geographically representative sample with flowers present on all images was selected.

All visible flowers in the 250 images were manually annotated using the CVAT tool (Sekachev et al., 2020) by drawing bounding boxes around the flowers. The dataset created has one class named 'flower'. This class includes all flowers with visible petals in the images, including flowers that are only partially within the image borders; the 'flower' category does not include grasses. To improve model performance and avoid errors due to large image size and high object density, each image was split into four slices. The final dataset follows the same point division as the initial dataset (60/20/20) and includes 300 image slices from 150 points for model training, 100 image slices from 50 points for validation data, 100 image slices from 50 points for test data. The slices coming from the same point were only ever used in one of the three datasets. Therefore, no image information is shared between training, validation, or test.

2.2 | Model training and parameter tuning

To detect individual flower in the images we trained the Faster R-CNN model on the manually delineated training dataset. Faster R-CNN is a computer vision model for object detection (Ren et al., 2015). This model improves the speed of the Fast R-CNN model by incorporating a region proposal network to extract initial object-like regions. The Faster R-CNN model was implemented using Detectron2, the Facebook AI Research library (Wu et al., 2019). We used the SGD optimizer with random clipping and Nesterov Accelerated Gradient following the original Faster R-CNN publication (Ren et al., 2015). This model has previously been used in similar workflows to estimate nectar mass from images (Hicks et al., 2021) and the extensions of this model was used to monitor phenology from time-lapse cameras (Mann et al., 2022).

We used weights pre-trained on the coco-dataset from the Detectron2 model zoo (Wu et al., 2019). Following preliminary tests we selected the ResNet-50 FPN model backbone with gamma of 0.1 and a ROI-heads batch size of 512. Finally, a hyper parameter tuning was done to explore two other model parameters: learning rate and momentum. 40 learning rates between 1e-8 and 1e-2, and 40 momentums between 0.9 and 0.99 were randomly generated. These intervals were selected following the values set in several other use cases using the same architecture (Hicks et al., 2021; Li et al., 2021; Ren et al., 2015; Sys et al., 2022). After 50 epochs half of the model settings were discarded and the 20 remaining models were trained for an additional 25 epochs.

2.2.1 | Performance metrics and final model selection

Performance metrics for the models are calculated based on the test dataset. The predicted flower is evaluated using an intersect

over union (IoU) threshold of 0.5. The IoU is equal to the area of the overlap (intersection) between the predicted bounding box and the reference bounding box divided by the area of their union. For each model all the objects of the test dataset were summed up into True positives (TP), False positives (FP), and False negatives (FN). Finally precision, recall, and F1, was calculated as follows Padilla et al. (2021):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F1\text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Performance metrics of the models were extracted three times during the workflow. The model with the highest F1 score was chosen as the best model. The F1 score provides a balanced score between precision and recall, thus is a measure of how precise and generalizing the model is. To evaluate the impact of different image input sizes on model performance, we further trained the final model on five datasets cropped to different sizes of 1500, 1000, 750, 500, and 224 pixels.

2.3 | Flower abundance, colour, and identification with Pl@ntNet

Having identified best performing model we predict the flower objects in the test dataset. The predicted flowers are then analysed to determine abundance and diversity of colour. The dominant colour pixel values is extracted from each predicted flower using a *k*-means clustering algorithm and the most dominant colour values are translated to a colour category. Finally, we attempt to identify the species of the predicted flowers using the Pl@ntNet application API (Affouard et al., 2017). Inspired by the work of August et al. (2020), a pipeline for querying each predicted flower individually with the Pl@ntNet API was developed. We set a threshold of 0.8 for the score of the identification and discuss how many flowers were 'correctly' identified using this methodology.

3 | RESULTS

3.1 | Dataset creation

The final 250 points selected for the train/validation/test dataset are mapped in Figure 3 with blue. As can be seen for the bar plots, the geographical distribution of the points in our sub-sample is proportional to the geographical distribution of the expert surveyed LUCAS grassland points. The lowest amount of points are selected for the Atlantic Northwest ($n=3$) grassland region and the

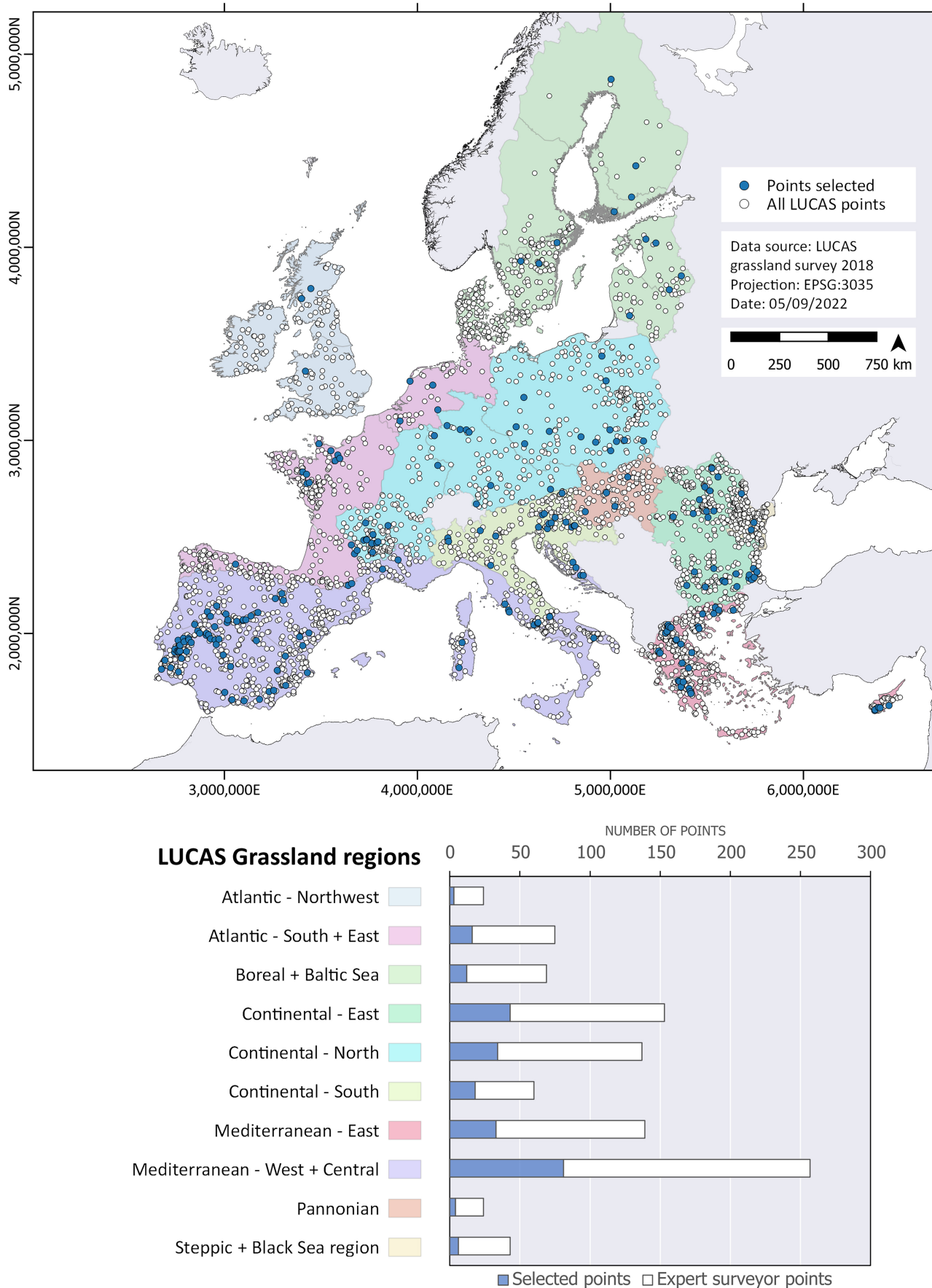


FIGURE 3 Map of the LUCAS grassland regions and all the surveyed points. Blue points are selected for the train/test data. The bar plot shows the number of points within each region for the expert surveyed points (white) from which our test/train dataset (blue) was sampled.

highest amount of points ($n=81$) are selected for the West Central Mediterranean grassland region.

Using CVAT we manually delineated 12,192 flowers on 250 images. The number of flowers delineated within one image ranges between 1 and 402 with an average of 59 flowers per image. Due to overpopulation in many of the images, each image was sliced in 4 slices and the delineations were cleaned, creating a final and improved version of the dataset consisting of 500 image slices with 9516 delineated flowers. The final dataset aimed to use two slices from each original image; however, since some images only had flowers present on one slice, 6% of the original images are represented by 1 slice each, 6% are represented by 3 slices each, and the remaining 88% are represented by 2 slices. The sliced images contain between 1 and 170 flowers each with an average of 25 per image slice. We use the native size of the images in our models. The image size ranged from 250×500 to 2000×2500 pixels. The majority of the images were of 1500×1800 pixels.

3.2 | Model tuning and performance

After an initial 50 epochs of training the first performance metrics were calculated and half of the hyper parameters defined in Section 2.2 were discarded. The remaining 20 models were trained an additional 25 epochs on the improved, sliced, and cleaned training dataset, and the second performance metrics were calculated, as summarized in Table 1. The best model was model setting 32 with a learning rate of $943e^{-6}$ and a momentum of 0.958. Based on the validation data, this model had the highest recall and F1 score compared to the other models, and thereby the best balance between precision and recall. The final performance metrics were extracted using the independent test set from ten runs of the best model. As seen in Table 2, the mean of the metrics are a precision of 0.89, recall of 0.61 and F1 score of 0.72. A full Table with all model parameters and performance metrics can be found in Supporting Information.

3.3 | Flower predictions

Using model setting 32 we inferred flowers on the test dataset. From the 100 images included in the test set 1377 flowers were detected, with a minimum of 1, and a maximum of 71. The colour of the

detected flowers was estimated using k -means colour clustering resulting in 849 yellow, 364 white, 87 purple, 30 blue, 9 pink, 7 grey, 9 red, 11 green, 6 brown and 5 orange flowers. Four visual examples of model outputs can be seen in Figure 4. In each image the boxes show the TP, FP, and FN. Finally the Table included in the image shows the flower abundance and colour distribution extracted from the predicted flowers of each image.

3.3.1 | Prediction errors

Visual inspection of the inference on the test dataset identified common errors summarized in Figure 5. From these examples we see that our results include a lot of 'hidden' flowers, for example objects that are flowers are classified as false positives in the model evaluation because they do not meet the IoU criteria. For example when the model and the reference flower unit do not match due to multiple overlapping flowers of the same species or when one flower is divided into multiple objects due to its inflorescence. This results in one or multiple reference objects within one prediction or vice versa, thus not passing the IoU requirements. Simply adjusting the IoU threshold did not increase model performance, as the problem is the variation in floral unit rather than the perfect delineation. Accounting for these 'hidden' TP requires a visual check of every single predicted object. However, to get an estimate of how much this affected the model performance metrics, hidden errors were 'caught' using the union between the reference and the prediction. If the area of the prediction is mostly within the reference or the other way around, then we count it as an overlap error and adjust this prediction to a TP. This rough estimate means that the cases where multiple reference flowers fit within one predicted flower or vice versa are adjusted to true flower detections and the performance metrics are recalculated. Including these 'hidden' TP increases the model recall on average by 12%.

3.4 | Pl@ntNet species detection

For each flower a crop of the original image was extracted using the predicted bounding box expanded with a buffer of 0.3 to ensure that the complete flower was within the image crop. The extracted image of the single flower was queried with the Pl@ntNet API. From the query results we extracted the best species identification and score

TABLE 1 First, second and adjusted performance metrics summarized by the min, median and max of all model performance metrics using the validation data.

	First metrics			Second metrics			Adjusted metrics		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Min	0.00	0.00	0.00	0.62	0.00	0.01	0.92	0.01	0.01
Med	0.80	0.07	0.13	0.89	0.55	0.68	0.95	0.61	0.75
Max	1.00	0.10	0.18	0.93	0.59	0.71	0.98	0.66	0.78

TABLE 2 Performance metrics extracted from 10 runs of the best model (32) based on the validation dataset and independent test dataset.

	Second metrics			Adjusted metrics		
	Precision	Recall	F1	Precision	Recall	F1
Best model (32) validation data						
Mean	0.89	0.58	0.70	0.95	0.65	0.77
Std	0.01	0.01	0.01	0.00	0.01	0.01
Best model (32) test data						
Mean	0.89	0.61	0.72	0.95	0.68	0.79
Std	0.01	0.01	0.01	0.01	0.01	0.01

from each predicted flower. As seen from [Figure 6](#), the first species probability score were generally low, ranging from 0.01 to 0.97, with a median of 0.1. This illustrates that we cannot always automatically determine the species of the predicted flowers using the available images. Only 6 out of the 1377 predicted flowers are above the rule of thumb score threshold recommended by Pl@ntNet of 0.8. Setting the threshold at 0.5 results in 52 predicted flowers. Expert botanist evaluation of those 52 images resulted in the identification at the species level of 22 flowers (and 33 flowers at the genus level), and confirmation of the predicted species in 16 cases (22 cases at the genus level). The botanist was not able to identify the species in 37 cases, and in 23 cases for the genus, mainly for the smallest image sizes. To summarize, 30% of the 52 flowers were correctly automatically identified at the species level, and 42% at the genus level.

4 | DISCUSSION

This study has shown that we can quantify flower abundance from an image with computer vision algorithms. Using the generic flower detection model developed in this paper we can capture individual flowers from an image. From the detected flowers we can extract various metrics such as abundances, colours, and sizes/shapes. The amount of floral resources and the diversity of flower shapes and colours in the ecosystem is important for pollinators (Trunschke et al., 2021). Using the flower detection algorithm makes it possible to quickly and efficiently monitor flower diversity and abundance by simply taking a photo of a patch of vegetation. Integrating this into large-scale monitoring schemes can decrease the surveyor bias and speed up the sampling process.

4.1 | Model performance and limitations

From [Tables 1](#) and [2](#) we can see how the two versions of our models and the adjusted IoU calculation, improved the performance metrics. Specifically interesting for this study is the recall, which is also where we see the largest improvement. As the aim is to develop a generic flower model capturing the large variation within the object category 'flower', the rate of omission is more important than the precision of the detection.

The flower predictions showed two main limitations of the detection algorithm: overpopulated images and complex scenes. The first issue was dealt with to a certain extent by slicing the initial images into four slices, thereby reducing the average number of objects in an image. This step had a large impact on the recall of the models, increasing the average recall by 49% from the first to the second performance metrics.

Since reducing image size to a fourth of the original had such a large impact we briefly explored reducing the size even further. From our final dataset we extracted several cropped dataset with a constant image size of 1500, 1000, 750, 500 and 224 pixels. We briefly trained the best model on each of the further cropped datasets. As seen from the performance metrics in [Table 3](#), smaller input size reduces precision and increases recall. This shows that even though we used quarter image slices of the original LUCAS grassland images, flower detection models can benefit from smaller and controlled image input size.

Apart from overpopulated images we also note the importance of image quality and a representative training dataset. The natural scenes captured in the images are complex. There are multiple flowers and plants overlapping, variations in image quality, light conditions, focus and, despite the protocol, taken at different distances to the plants. These complex scenes are reflected in the imperfect delineations of the train/test data. Although the delineations were corrected several times, there are still tiny flowers omitted by accident, areas where multiple flowers overlapping are delineated as one single flower, and variations in the definition of the floral units due to different inflorescence of the flowers and different growth stages. For example, from a computer vision perspective flowers from the umbellifers family (i.e. Apiaceae) with an umbel inflorescence can appear as a single flower unit or several separate flower units depending on the growth stage/density of the umbel. These variations in flower unit and omissions of small or unclear flowers in the delineated flowers in the result, has for consequence the fact that many of the prediction errors being 'hidden flowers', for example objects that are classified as false positives even though they are actually flowers, as seen in in [Figure 5](#). To briefly explore how much such errors influenced the model performance we reclassified all predicted objects within the reference flower object and vice versa into true positives and calculate the adjusted performance metrics. With this rough estimate we increased the model recall with 11% with respect to the second set of performance metrics. However, in order to understand the full impact of such errors in this work visual, inspection of all predictions is needed.

To mitigate the issues with imperfect delineations and floral unit variation, further development of the training dataset is needed for better representation of the huge variation in the floral domain. Both capturing the large amount of flower species that could occur but also capturing the within species variation resulting from growth stage and regional variability. Advances are being made in this field, focusing on two species, Mann et al. (2022) quantified fine-scale flower abundance and phenology dynamics on images using deep learning. Finally, improved deep learning models may also be tested. For example vision transformers trained in a self-supervised way

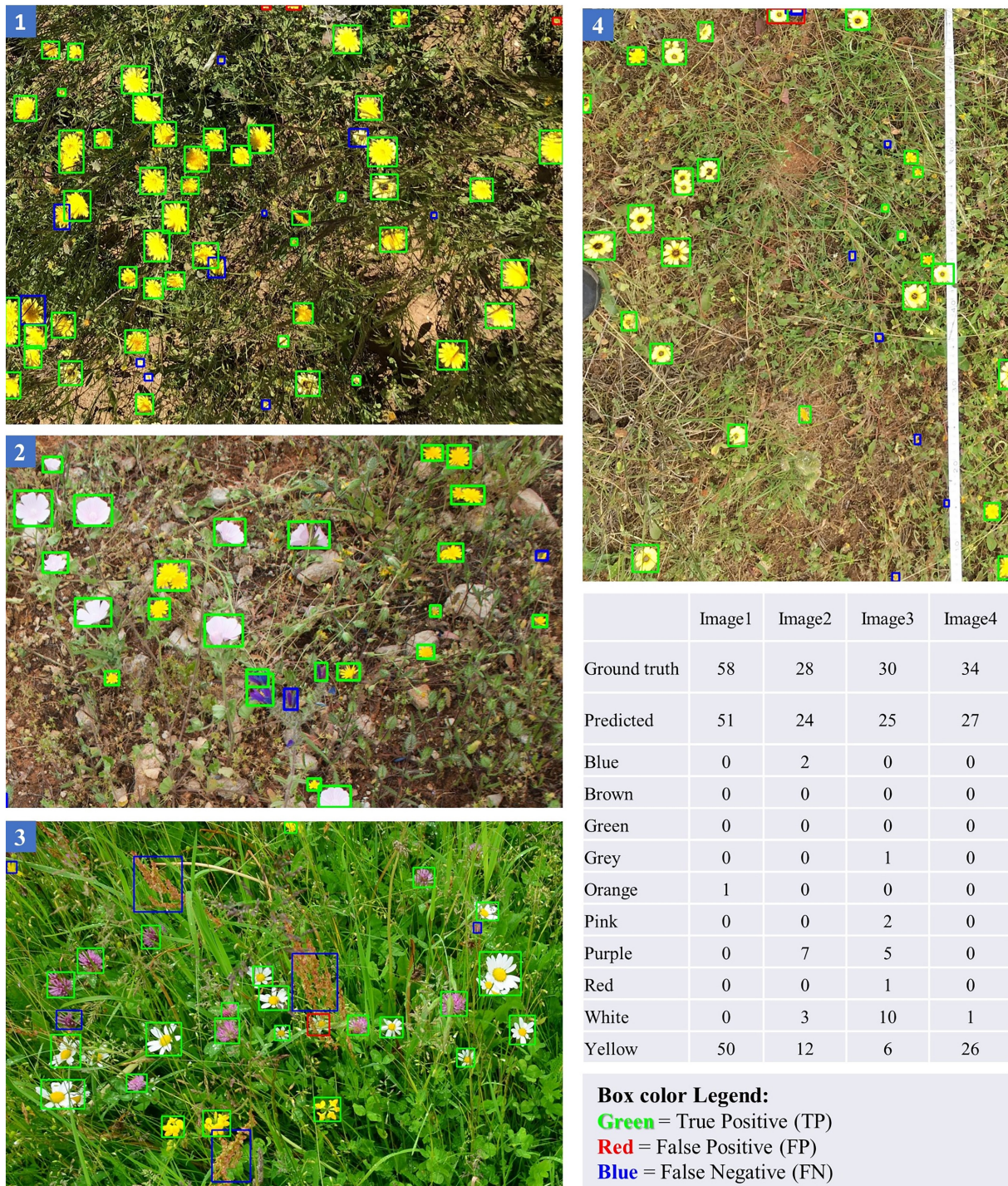


FIGURE 4 Example of inference using the best model on four images from the test dataset. The table shows the final counts of the reference delineated flowers and predicted flowers in each image as well as predicted colour.

that have been shown to have better segmentation capabilities than CNNs (Bao et al., 2021; Caron et al., 2021).

4.2 | Species identification

Pl@ntNet queries of the predicted flowers in the test dataset generally resulted in low identification scores. From Figure 6 we see

that the few flowers with higher identification scores are larger and brighter images. This indicates that the image quality of the individual flower crops are not sufficient for this recognition algorithm, because the flower images extracted from the LUCAS grassland data do not correspond to the type of imagery used to train the algorithm. Pl@ntNet is primarily trained on closeup and sharp images focusing on a single plant, while the images from the LUCAS grassland survey show a patch of vegetation. Previous studies shows that the input

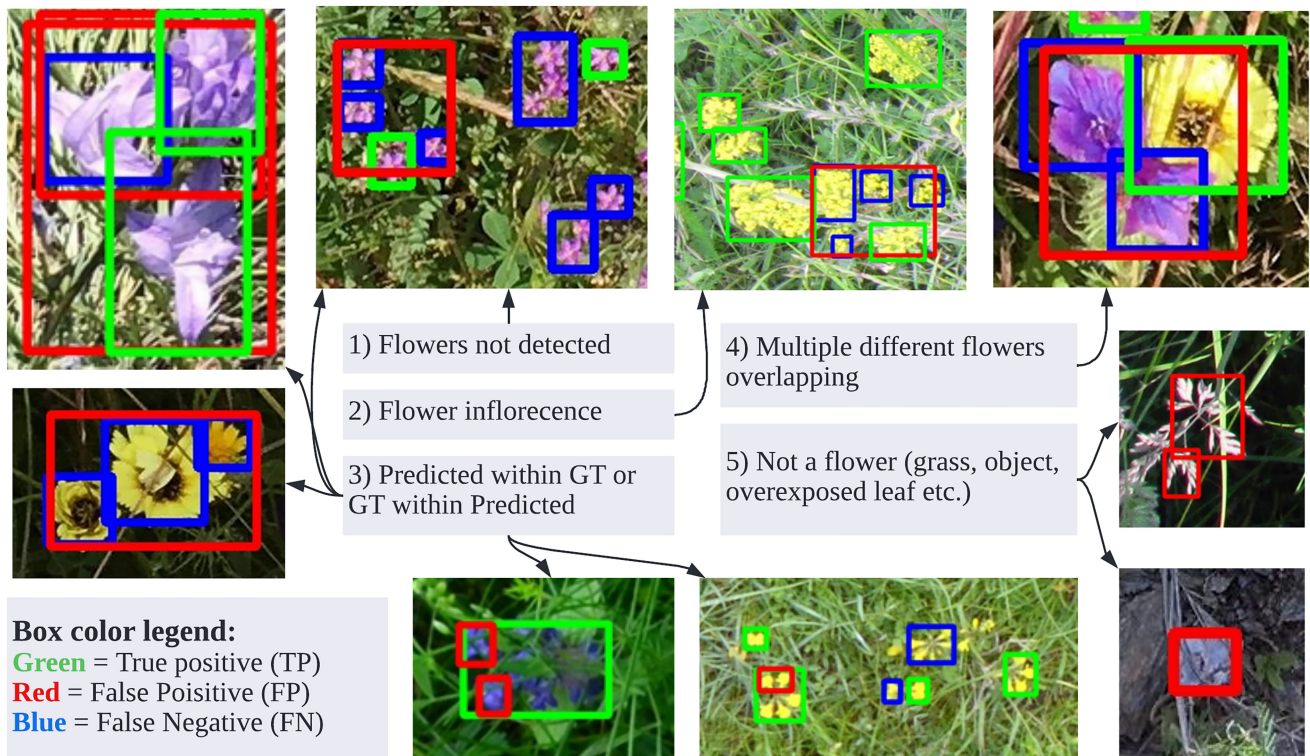


FIGURE 5 Closeup visual examples of some common errors occurring in the inference of the test dataset.

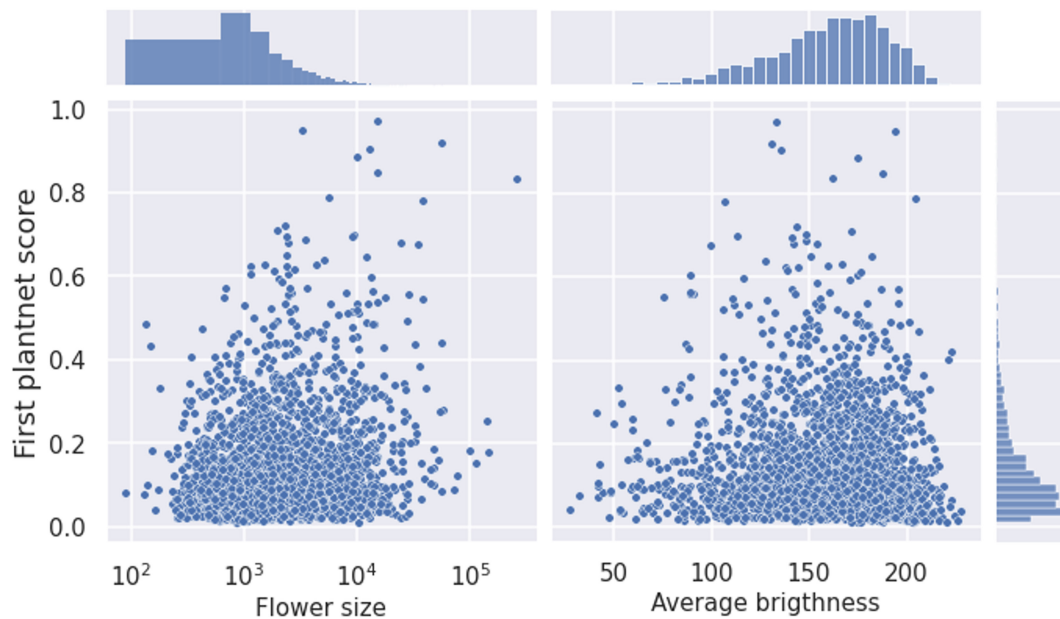


FIGURE 6 The top prediction score from the PI@ntNet species identification queries of all the predicted flowers in the test dataset. The scatter plot shows the distribution of scores according to size (left) and average brightness, for example average grey scale pixel values (right).

image quality is vital for species identification algorithms (Rzanny et al., 2019; Wäldchen et al., 2018). Furthermore, using several pictures of various flower organs and ancillary information increases the identification certainty. The imagery provided by the LUCAS grassland module was not taken with computer vision purposes in mind, therefore, does not include closeup images of high quality that

the identification requires. Image crops of individual flowers are in this case sufficient for detecting colour groups but not for identification of individual species at large scale. Perhaps identification at a coarser taxonomic rank would result in higher confidence score and thereby reliable information on flower taxa could still be extracted from images of vegetation patches.

TABLE 3 Best model trained on further cropped flower datasets with fixed input image size. Performance metrics extracted based on the independent test set.

	Precision	Recall	F1 score
Original data	0.88	0.62	0.73
Crop 224	0.18	0.77	0.29
Crop 500	0.52	0.83	0.64
Crop 750	0.70	0.77	0.73
Crop 1000	0.82	0.72	0.76
Crop 1500	0.90	0.62	0.74

Improving image input through upscaling the cropped flowers is another option to enhance the identification. This analysis found that simply expanding the predicted bounding box by 30%, and thereby including some background in the images to be identified resulted in better identification scores from Pl@ntNet. Finally, previous studies have showed that including ancillary information such as geo-location could improve identification scores.

4.3 | Outlook

The 2018 LUCAS grassland survey is just one example of a large-scale biodiversity monitoring scheme. Large-scale surveys are the outcome of an optimisation procedure, aiming at gathering the maximum of information while containing the costs. The time the surveyor needs to remain on the point, to be able to survey all requested parameters, is, therefore, a key element of a survey. Incorporating computer vision into such workflows show a great potential for extracting valuable information on biodiversity using quick snapshots taken in the field, and increasing the accuracy of the output for selected parameters. Such automated methods can also support the process of CAP compliance assurance based on pictures provided by farmers. In fact, in Thuringen, Germany, farmers can make use of the Flora Incognita app (Mäder et al., 2021) to provide evidence on their management of environmentally sensitive grasslands.

Further improvement of the detection workflow could include benchmarking other backbone options for Faster R-CNN like Transformer-based backbones (Dian et al., 2022), or different models like Detection with Multi-modal Transformer (Maaz et al., 2021) or Hierarchical Shot Detector (Cao et al., 2019). The last two models mentioned show high MAP (mean average precision) in the PASCAL VOC (2007) benchmarking. However, taking into account that flower detection is a binary classification, the improvements in precision probably will not be proportional to the increase in model complexity. This can be seen from the benchmark with the Oxford flower dataset (Flowers-102 Benchmark, n.d.). Even though this benchmark is based on image classification and not object detection, we can see that the introduction of various SOTA models did not lead to significant performance improvements. Whether the inherent simplicity of our specific flower detection task may limit the impact of adopting

more complex models needs to be investigated further in future work. Nonetheless, for our methodological workflow, the choice of Faster R-CNN with a ResNet-50 backbone remains a good option.

Improvements to structured surveys with automated workflows is one of the ingredients to improve overall biodiversity monitoring at scale. Increasingly information is also gathered ad-hoc through citizen science activities. Explicitly integrating the sampling design of, for example the LUCAS grassland biodiversity or EMBAL (Environmental Monitoring of Biodiversity in Agricultural Landscapes) surveys in citizen science apps could improve the temporal sampling at those points. By using the algorithm we developed, a single picture of the sampling point can generate information on floral diversity and abundance in a systematic way. Besides the single species recognition capacity of current apps, surveyors and citizen scientists will benefit from instantaneously obtaining community level information on floral diversity. In the future, this will likely expand to the identification of multiple species or species assemblages. In this study we focused on identifying multiple flower objects on a single photo, species identification on these objects proved difficult. Future endeavours in this field may quantify multiple species on a single photo instantaneously using advanced deep learning approaches. Platforms such as GBIF are bringing all these datastreams together. A possible next step is to create feedback loops from such repositories and link them to the type of photos that we have used here and that contain a mix of flowers. This could allow upscaling to computer vision models that directly derive habitat related information from such imagery.

The LUCAS grassland survey was repeated in 2022 with 20K points with images surveyed throughout European grasslands; simultaneously image data is increasingly generated through citizen science projects. Exploiting this large amount of images through computer vision can give us an understanding of biodiversity covering temporal and spatial scales not possible through traditional surveying methods. With more focus on how to extract meaningful information from these images of varying quality we can unlock information needed for better targeting conservation and restoration efforts in agricultural landscapes and other types of landscapes with floral resources, and to support specific initiatives on pollinators (Duque-Trujillo et al., 2022; European Commission Eu Pollinators Initiative, 2018). The generic flower detection algorithm developed in this research is a step in this direction. The algorithm will be enriched to improve generalization ability, and in the future may be included in operational services such as Pl@ntNet with post-processing results tailored to particular objectives (e.g. quantifying colour diversity).

5 | CONCLUSIONS

A generic flower object detector was built using Faster R-CNN trained on 2018 LUCAS surveyed top-down looking grassland imagery. Individual flowers were successfully detected with a precision and recall of, respectively, 0.89 and 0.61 Biodiversity relevant

data on abundance and colour distribution were automatically retrieved. Due to the image quality of extracted flower objects, species identification with Pl@ntNet was a challenge. The research presented here allows for a further integration of computer vision-based tools in large-scale biodiversity monitoring of grassland flowering plants.

AUTHOR CONTRIBUTIONS

M. van der Velde, M. L. Paracchini, L. Martinez-Sanchez and N. Elvekjaer conceived the ideas and designed methodology; M. L. Paracchini, N. Elvekjaer and L. Martinez-Sanchez collected the data; N. Elvekjaer, L. Martinez-Sanchez, P. Bonnet, A. Joly analysed the data; N. Elvekjaer and M. van der Velde led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of this research by the JRC Exploratory Research Program through the Rural Refocus Project (31280). This work was partially supported by the European Union's Horizon 2020 Research and Innovation Program under grant agreement no 863463 (Cos4Cloud project). This work benefited from services and resources provided by the EGI-ACE project (receiving funding from the European Union's Horizon 2020 Research and Innovation under grant agreement no. 101017567), with the dedicated support of IFCA.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/2688-8319.12324>.

DATA AVAILABILITY STATEMENT

The dataset created in this study including the images and annotations formatted in the coco annotation format along with relevant LUCAS data is available via the JRC data catalogue (European Commission, Joint Research Centre, 2022): <https://data.jrc.ec.europa.eu/dataset/caa582b7-7f45-4748-9223-08e5f145a4a6>.

ORCID

N. Elvekjaer  <https://orcid.org/0000-0002-7655-9287>

L. Martinez-Sanchez  <https://orcid.org/0000-0002-1354-1792>

P. Bonnet  <https://orcid.org/0000-0002-2828-4389>

A. Joly  <https://orcid.org/0000-0002-2161-9940>

M. van der Velde  <https://orcid.org/0000-0002-9103-7081>

REFERENCES

Affouard, A., Goëau, H., Bonnet, P., Lombardo, J. C., Affouard, A., Goëau, H., Bonnet, P., Lombardo, J. C., Joly, A., Christophe Lombardo, J., & Joly, A. (2017). ICLR (*International Conference On Learning*

Representations), 2017 Workshop. Pl@ntNet App in the Era of Deep Learning.

August, T., Harvey, M., Lightfoot, P., Kilbey, D., Papadopoulos, T., & Jepson, P. (2015). Emerging technologies for biological recording. *Biological Journal of the Linnean Society*, 115, 731–749. <https://doi.org/10.1111/BJ.12534>

August, T., Pescott, O., Joly, A., & Bonnet, P. (2020). Ai naturalists might hold the key to unlocking biodiversity data in social media imagery. *Patterns*, 1, 100116. <https://doi.org/10.1016/j.patter.2020.100116>

Bao, H., Dong, L., Piao, S., & Wei, F. (2021). Beit: Bert pre-training of image transformers. <https://doi.org/10.48550/arxiv.2106.08254>

Barmeier, G. (2021). Detection of individual plants on grassland by using geotagged photos. https://marswiki.jrc.ec.europa.eu/wikicap/images/5/57/09_Detection_of_individual_plants.pdf

Bartual, A. M., Sutter, L., Bocci, G., Moonen, A. C., Cresswell, J., Entling, M., Giffard, B., Jacot, K., Jeanneret, P., Holland, J., Pfister, S., Pintér, O., Veromann, E., Winkler, K., & Albrecht, M. (2019). The potential of different semi-natural habitats to sustain pollinators and natural enemies in european agricultural landscapes. *Agriculture, Ecosystems and Environment*, 279, 43–52. <https://doi.org/10.1016/J.AGEE.2019.04.009>

Boetzl, F. A., Krauss, J., Heinze, J., Hoffmann, H., Juffa, J., König, S., Krimmer, E., Prante, M., Martin, E. A., Holzschuh, A., & Steffan-Dewenter, I. (2021). A multitaxa assessment of the effectiveness of agri-environmental schemes for biodiversity management. *Proceedings of the National Academy of Sciences of the United States of America*, 118, e2016038118. <https://doi.org/10.1073/pnas.2016038118>

Boho, D., Rzanny, M., Wäldchen, J., Nitsche, F., Deggelmann, A., Wittich, H. C., Seeland, M., & Mäder, P. (2020). Flora capture: A citizen science application for collecting structured plant observations. *BMC Bioinformatics*, 21, 576. <https://doi.org/10.1186/S12859-020-03920-9>

Bonnet, P., Joly, A., Faton, J., Brown, S., Kimiti, D., Deneu, B., Servajean, M., Affouard, A., Lombardo, J., Mary, L., Vignau, C., & Munoz, F. (2020). How citizen scientists contribute to monitor protected areas thanks to automatic plant identification tools. *Ecological Solutions and Evidence*, 1, e12023. <https://doi.org/10.1002/2688-8319.12023>

Bonnet, P., Joly, A., Goëau, H., Champ, J., Vignau, C., Molino, J.-F., Barthélémy, D., & Boujema, N. (2016). Plant identification: Man vs. machine. *Multimedia Tools and Applications*, 75, 1647–1665.

Botella, C., Joly, A., Bonnet, P., Monestiez, P., & Munoz, F. (2018). Species distribution modeling based on the automated identification of citizen observations. *Applications in Plant Sciences*, 6, 1–11. <https://doi.org/10.1002/aps.3.1029>

Cao, J., Pang, Y., Han, J., & Li, X. (2019). Hierarchical shot detector. In *Proceedings of the IEEE International Conference on Computer Vision Foundation*, 2019-October, pp. 9704–9713.

Caron, M., Touvron, H., Misra, I., Jegou, H., Mairal, J., Bojanowski, P., & Joulin, A. (2021). Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE International Conference on Computer Vision Foundation*. <https://doi.org/10.48550/arxiv.2104.14294>

Cole, L. J., Kleijn, D., Dicks, L. V., Stout, J. C., Potts, S. G., Albrecht, M., Balzan, M. V., Bartomeus, I., Bebeli, P. J., Bevk, D., Biesmeijer, J. C., Chlebo, R., Dautarté, A., Emmanouil, N., Hartfield, C., Holland, J. M., Holzschuh, A., Knoben, N. T. J., Kovács-Hostyánszki, A., ... Scheper, J. (2020). A critical analysis of the potential for eu common agricultural policy measures to support wild pollinators on farmland. *Journal of Applied Ecology*, 57(4), 681–694. <https://doi.org/10.1111/1365-2664.13572>

d'Andrimont, R., Yordanov, M., Martinez-Sanchez, L., Eiselt, B., Palmieri, A., Dominici, P., Gallego, J., Reuter, H. I., Joebges, C., Lemoine, G., & van der Velde, M. (2020). Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European

- Union. *Scientific Data*, 7, 352. <https://doi.org/10.1038/S41597-020-00675-Z>
- de Lutio, R., She, Y., D'Aronco, S., Russo, S., Brun, P., Wegner, J. D., & Schindler, K. (2021). Digital taxonomist: Identifying plant species in community scientists' photographs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 182, 112–121. <https://doi.org/10.1016/J.ISPRSJPRS.2021.10.002>
- Dian, S., Zhong, X., & Zhong, Y. (2022). Faster r-transformer: An efficient method for insulator detection in complex aerial environments. *Measurement*, 199, 111238.
- Duque-Trujillo, D., Hincapié, C. A., Osorio, M., & Zartha-Sossa, J. W. (2022). Strategies for the attraction and conservation of natural pollinators in agroecosystems: A systematic review. *International Journal of Environmental Science and Technology*, 20, 1–14. <https://doi.org/10.1007/S13762-022-04634-6/TABLES/2>
- European Commission Eu Pollinators Initiative. (2018). <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528213737113&uri=CELEX:52018D0395>
- European Commission, Joint Research Centre. (2022). Flower detection dataset PID. <http://data.europa.eu/89h/caa582b7-7f45-4748-9223-08e5f145a4a6>
- Flowers-102 Benchmark. (n.d.). <https://paperswithcode.com/sota/image-classification-on-flowers-102>
- Hicks, D., Baude, M., Kratz, C., Ouvrard, P., & Stone, G. (2021). Deep learning object detection to estimate the nectar sugar mass of flowering vegetation. *Ecological Solutions and Evidence*, 2, e12099. <https://doi.org/10.1002/2688-8319.12099>
- Høyev, T. T., Årje, J., Bjerger, K., Hansen, O. L., Iosifidis, A., Leese, F., Mann, H. M., Meissner, K., Melvad, C., & Raitoharju, J. (2021). Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences of the United States of America*, 118(2), e2002545117. <https://doi.org/10.1073/PNAS.2002545117>
- IPBES. (2019). Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. <https://zenodo.org/record/6417333>
- Joly, A., Goëau, H., Bonnet, P., Bakić, V., Barbe, J., Selmi, S., Yahiaoui, I., Carré, J., Mouysset, E., Molino, J. F., Boujemaa, N., & Barthélémy, D. (2014). Interactive plant identification based on social image data. *Ecological Informatics*, 23, 22–34. <https://doi.org/10.1016/j.ecoinf.2013.07.006>
- Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.-P., Lombardo, J.-C., Planqué, R., Palazzo, S., & Müller, H. (2019). Biodiversity information retrieval through large scale content-based identification: A long-term evaluation. In *Information retrieval evaluation in a changing world* (pp. 389–413). Springer Nature. https://doi.org/10.1007/978-3-030-22948-1_16
- Jones, H. G. (2020). What plant is that? Tests of automated image recognition apps for plant identification on plants from the british flora. *AoB Plants*, 12, plaa052. <https://doi.org/10.1093/AOBPLA/PLAA052>
- Li, Z., Li, Y., Yang, Y., Guo, R., Yang, J., Yue, J., & Wang, Y. (2021). A high-precision detection method of hydroponic lettuce seedlings status based on improved Faster RCNN. *Computers and Electronics in Agriculture*, 182, 106054. <https://doi.org/10.1016/j.compag.2021.106054>
- Maaz, M., Rasheed, H., Khan, S., Khan, F. S., Anwer, R. M., & Yang, M. H. (2021). Class-agnostic object detection with multi-modal transformer. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13670 LNCS:512–531.
- Mäder, P., Boho, D., Rzanny, M., Seeland, M., Wittich, H. C., Deggelmann, A., & Wäldchen, J. (2021). The flora incognita app—Interactive plant species identification. *Methods in Ecology and Evolution*, 12, 1335–1342. <https://doi.org/10.1111/2041-210X.13611>
- Mahecha, M. D., Rzanny, M., Kraemer, G., Mäder, P., Seeland, M., & Wäldchen, J. (2021). Crowd-sourced plant occurrence data provide a reliable description of macroecological gradients. *Ecography*, 44, 1131–1142. <https://doi.org/10.1111/ecog.05492>
- Mann, H. M. R., Iosifidis, A., Jepsen, J. U., Welker, J. M., Loonen, M. J. J. E., & Høyev, T. T. (2022). Automatic flower detection and phenology monitoring using time-lapse cameras and deep learning. *Remote Sensing in Ecology and Conservation*, 8, 765–777. <https://doi.org/10.1002/rse2.275>
- Nilsson, L., Klatt, B. K., & Smith, H. G. (2021). Effects of flower-enriched ecological focus areas on functional diversity across scales. *Frontiers in Ecology and Evolution*, 9, 629124. <https://doi.org/10.3389/FEVO.2021.629124>
- Oppermann, R. (2021). Lucas grassland module 2018—Description and validation. <https://ec.europa.eu/eurostat/documents/205002/9722562/Grassland-Modelling-Report.pdf>
- Padilla, R., Passos, W. L., Dias, T. L., Netto, S. L., & Silva, E. A. D. (2021). A comparative analysis of object detection metrics with a companion open-source toolkit. *Electronics*, 10, 279. <https://doi.org/10.3390/ELECTRONICS10030279>
- PASCAL VOC. (2007). Benchmark. <https://paperswithcode.com/sota/object-detection-on-pascal-voc-2007>
- Pl@ntNet Contribution to GBIF (Global Biodiversity Information Facility). (2024). <https://www.gbif.org/publisher/da86174a-a605-43a4-a5e8-53d484152cd3>
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- Rzanny, M., Mäder, P., Deggelmann, A., Chen, M., & Wäldchen, J. (2019). Flowers, leaves or both? How to obtain suitable images for automated plant identification. *Plant Methods*, 15(1), 77. <https://doi.org/10.1186/s13007-019-0462-4>
- Schaub, S., Finger, R., Leiber, F., Probst, S., Kreuzer, M., Weigelt, A., Buchmann, N., & Scherer-Lorenzen, M. (2020). Plant diversity effects on forage quality, yield and revenues of semi-natural grasslands. *Nature Communications*, 11, 768. <https://doi.org/10.1038/s41467-020-14541-4>
- Scheper, J., Bommarco, R., Holzschuh, A., Potts, S. G., Riedinger, V., Roberts, S. P. M., Rundlöf, M., Smith, H. G., Steffan-Dewenter, I., Wickens, J. B., Wickens, V. J., & Kleijn, D. (2015). Local and landscape-level floral resources explain effects of wildflower strips on wild bees across four European countries. *Journal of Applied Ecology*, 52, 1165–1175. <https://doi.org/10.1111/1365-2664.12479>
- Scheper, J., Holzschuh, A., Kuussaari, M., Potts, S. G., Rundlöf, M., Smith, H. G., & Kleijn, D. (2013). Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss—A meta-analysis. *Ecology Letters*, 16, 912–920. <https://doi.org/10.1111/ELE.12128>
- Scholes, R. J., Walters, M., Turak, E., Saarenmaa, H., Heip, C. H., Tuama, É. Ó., Faith, D. P., Mooney, H. A., Ferrier, S., Jongman, R. H., Harrison, I. J., Yahara, T., Pereira, H. M., Larigauderie, A., & Geller, G. (2012). Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, 4, 139–146. <https://doi.org/10.1016/J.COSUST.2011.12.005>
- Seeland, M., Rzanny, M., Boho, D., Wäldchen, J., & Mäder, P. (2019). Image-based classification of plant genus and family for trained and untrained plant species. *BMC Bioinformatics*, 20, 1–14. <https://doi.org/10.1186/s12859-018-2474-x>
- Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., Tosmanov, Kruchinin, D., Zankevich, A., DmitriySidnev, M. M., Johannes222, Chenuet, M., a andre, telenachos, Melnikov, A., Kim, J., Ilouz, L., Glazov, N., Priya4607, Tehrani, R., ... Truong, T. (2020). Computer vision annotation tool (cvat). <https://zenodo.org/record/4009388>

- Sima, A., Loudjani, P., & Devos, W. (2020). Use of geotagged photographs in the frame of common agriculture policy checks. https://marsw.iki.jrc.ec.europa.eu/wikicap/images/c/ce/Geotagged_JRC_Report1.pdf
- Sutcliffe, L., Schraml, A., Eiselt, B., & Oppermann, R. (2019). The LUCAS grassland module pilot—Qualitative monitoring of grassland in Europe. *Palaearctic Grasslands*, 40, 27–31. <https://doi.org/10.21570/EDGG.PG.40.27-31>
- Sutter, L., Jeanneret, P., Bartual, A. M., Bocci, G., & Albrecht, M. (2017). Enhancing plant diversity in agricultural landscapes promotes both rare bees and dominant crop-pollinating bees through complementary increase in key floral resources. *Journal of Applied Ecology*, 54, 1856–1864. <https://doi.org/10.1111/1365-2664.12907>
- Sys, S., Weißbach, S., Jakob, L., Gerber, S., & Schneider, C. (2022). Collembolai, a macrophotography and computer vision workflow to digitize and characterize samples of soil invertebrate communities preserved in fluid. *Methods in Ecology and Evolution*, 13(12), 2729–2742. <https://doi.org/10.1111/2041-210X.14001>
- Terry, J. C. D., Roy, H. E., & August, T. A. (2020). Thinking like a naturalist: Enhancing computer vision of citizen science images by harnessing contextual data. *Methods in Ecology and Evolution*, 11, 303–315. <https://doi.org/10.1111/2041-210X.13335>
- Trunschke, J., Lunau, K., Pyke, G. H., Ren, Z. X., & Wang, H. (2021). Flower color evolution and the evidence of pollinator-mediated selection. *Frontiers in Plant Science*, 12, 1096. <https://doi.org/10.3389/FPLS.2021.617851/BIBTEX>
- Unger, S., Rollins, M., Tietz, A., & Dumais, H. (2021). Inaturalist as an engaging tool for identifying organisms in outdoor activities. *Journal of Biological Education*, 55, 537–547. <https://doi.org/10.1080/00219266.2020.1739114>
- Villon, S., Mouillot, D., Chaumont, M., Subsol, G., Claverie, T., & Villéger, S. (2020). A new method to control error rates in automated species identification with deep learning algorithms. *Scientific Reports*, 10, 12. <https://doi.org/10.1038/S41598-020-67573-7>
- Wäldchen, J., Rzanny, M., Seeland, M., & Mäder, P. (2018). Automated plant species identification—Trends and future directions. *PLoS Computational Biology*, 14, 1–20. <https://doi.org/10.1371/journal.pcbi.1005993>
- Weinstein, B. G. (2018). A computer vision for animal ecology. *Journal of Animal Ecology*, 87(3), 533–545. <https://doi.org/10.1111/1365-2656.12780>
- Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., & Girshick, R. (2019). Detectron2. <https://github.com/facebookresearch/detectron2>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Table S1. Model parameters used to find the optimal learning rate and momentum.

Table S2. Performance metrics extracted from the training and validation data.

How to cite this article: Elvekjaer, N., Martinez-Sanchez, L., Bonnet, P., Joly, A., Paracchini, M. L., & van der Velde, M. (2024). Detecting flowers on imagery with computer vision to improve continental scale grassland biodiversity surveying. *Ecological Solutions and Evidence*, 5, e12324. <https://doi.org/10.1002/2688-8319.12324>