

RESEARCH ARTICLE

Deep learning object detection to estimate the nectar sugar mass of flowering vegetation

Damien Hicks¹  | Mathilde Baude² | Christoph Kratz¹ | Pierre Ouvrard³  |
Graham Stone⁴

¹ Statistics and Modelling Service, Foss House, Kings Pool, Natural England, York, UK

² LBLGC, Equipe entomologie et biologie intégrée, Rue de Chartres, BP 6759, Université d'Orléans, Orleans, France

³ Osmia, Technopole Agropole, Estillac, France

⁴ Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

Correspondence

Damien Hicks, Statistics and Modelling Service, Natural England, Foss House, Kings Pool, 1–2 Peasholme Green, New York, YO1 7PX, UK.
Email: damien.hicks@naturalengland.org.uk

Funding information

Natural England; POLLEN; Microsoft AI for Earth

Handling Editor: Marc Cadotte

[Correction added on 28 September 2021 after first online publication: affiliation of Christoph Kratz is updated.]

Abstract

1. Floral resources are a key driver of pollinator abundance and diversity, yet their quantification in the field and laboratory is laborious and requires specialist skills.
2. Using a dataset of 25,000 labelled tags of fieldwork-realistic quality, a convolutional neural network (Faster R-CNN) was trained to detect the nectar-producing floral units of 25 taxa in surveyors' quadrat images of native, weed-rich grassland in the United Kingdom.
3. Floral unit detection on a test set of 50 model-unseen images of comparable vegetation returned a precision of 90%, recall of 86% and F1 score (the harmonic mean of precision and recall) of 88%. Model performance was consistent across the range of floral abundance in this habitat.
4. Comparison of the nectar sugar mass estimates made by the CNN and three human surveyors returned similar means and standard deviations. Over half of the nectar sugar mass estimates made by the model fell within the absolute range of those of the human surveyors.
5. The optimal number of quadrat image samples was determined to be the same for the CNN as for the average human surveyor. For a standard quadrat sampling protocol of 10–15 replicates, this application of deep learning could cut pollinator-plant survey time per stand of vegetation from hours to minutes.
6. The CNN is restricted to a single view of a quadrat, with no scope for manual examination or specimen collection, though in contrast to human surveyors its object detection is deterministic and its floral unit definition is standardized.
7. As agri-environment schemes move from prescriptive to results-based, this approach provides an independent barometer of grassland management which is usable by both landowner and scheme administrator. The model can be adapted to visual estimations of other ecological resources such as winter bird food, floral pollen volume, insect infestation and tree flowering/fruitletting, and by adjustment of classification threshold may show acceptable taxonomic differentiation for presence–absence surveys.

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.

© 2021 The Authors. *Ecological Solutions and Evidence* published by John Wiley & Sons Ltd on behalf of British Ecological Society

KEYWORDS

convolutional neural network, deep learning, Faster R-CNN, nectar sugar mass, object detection, pollinator food resources, vegetation survey

1 | INTRODUCTION

Reduction of floral resources, primarily through habitat loss and changes in land use, is indicated as a major cause of insect pollinator decline (Dicks et al., 2015; Jones et al., 2021). Intervention strategies focus on enhancing pollinator food resources by measures such as crop diversification, relaxed mowing (e.g. Phillips et al., 2020), sown flower meadows (Ouvrard et al., 2018) and crop margins (Carvell et al., 2004).

Nectar and pollen-rich habitat are key drivers of pollinator abundance and diversity (Baldock et al., 2015; Timberlake et al. 2021), yet their analysis in the field and laboratory is laborious and requires specialist skills (Breeze et al., 2020). The standard field survey method to monitor and quantify pollinator food plants (comprising: species identification, flower counts in replicated quadrats/transects, then conversion to pollinator resource) requires a prohibitively large amount of person-power to obtain a robust sample; for example Baldock et al. (2019) employed a team of 12 people to count two million flowers over 2 years. Field images of quadrats are usually taken for confirmatory purposes only, by necessity recording a general overview rather than focused portraits of the individual plant species present.

The aim of this work was to determine the potential for automated estimation of the daily nectar sugar mass produced by stands of flowering vegetation.

1.1 | Computer vision

Automated classification has been successfully applied to identify single instances of taxa in images (e.g. Tabak et al., 2019; Valan et al., 2019). Our challenge was to detect multiple instances of floral units per image, of a range of taxa in vegetation quadrats. This requires (i) generating multiple bounding boxes ('tags') surrounding candidate objects (floral units) in a given image, (ii) labelling each object as belonging to a specific class (here, 25 plant taxa) and (iii) assigning a confidence score to each object classification.

The raw data are digital colour images, typically consisting of three matrices (the 'channels' red, green, and blue) each with a total number of pixels equalling $image\ height \times image\ width$. The individual pixel values range from 0 to 255, so each colour in the image is one of (256^3) 16.8 million possibilities.

1.2 | Deep learning

The term 'deep' defines a subfield of machine learning characterized by hierarchical representations (Bengio, 2012), which has recently achieved historic breakthroughs in image classification, speech recog-

niton, text-to-speech conversion and board games, by using neural networks to learn features in data. Such models often use hundreds of successive layers of representations learned automatically from training datasets, the retained representations becoming increasingly abstract from the original data and increasingly informative about the target. The learning in such a network therefore is an optimization process in which each layer is adjusted for optimal performance on the training data via a feedback signal (the loss function).

Deep learning is within reach of ecologists thanks to powerful open-source software and the availability of graphical processing units (GPUs) as hardware or virtual machines. For labour-intensive ecological surveys, Torney et al. (2019) concluded that their deep learning network of Wildebeest abundance could reduce 18 person-weeks of analysis to 1 day, to within 1% of the estimates of expert human surveyors.

1.3 | Convolutional neural networks

Convolutional neural networks (CNN) are currently one of the most promising deep learning methods for ecological studies. CNNs are designed to learn feature hierarchies and local patterns in a stochastic way, typically learning these in small two-dimensional frames of input images. The building blocks of a CNN are data-processing modules (the convolutional layers), which by matrix arithmetic act as filters, extracting representations from the data to produce a trained network. The numerical arrays of pixel values in the training dataset are inputted to the CNN, then its predictions are iteratively compared to the labelled validation data by backpropagation (see below). As the CNN has not been exposed to the validation data, its performance on this dataset indicates whether the model is underfitting or overfitting, that is how well it is generalizing to unseen instances.

The loss function guides an iterative search for weight values to map inputs to outputs (in this case, floral units) by successively adjusting the value of the weights in a direction that lowers the loss score, then repeating over many training-loops (epochs). Each labelled box in the training image set and labelled box proposed by the model are thus evaluated to estimate accuracy. As the matrix operations of the CNN are differentiable, the gradient of the loss can be computed and weights moved in the opposite direction from the gradient, so to minimize the loss function and yield outputs which are as close as possible to inputs. Backpropagation applies the chain rule of calculus to compute the effect that each parameter had in the validation loss, and the updated parameters are used to start the next epoch. A multiplicative factor specified by the modeller (the learning rate) is applied to the gradient for the parameter update. An epoch is complete when every bounding box has already been used for computing the loss.

In early work, Chen et al. (2014) used CNNs for animal identification to a detection accuracy of 0.38 using a dataset (20,000 images and 20 classes) of a similar scale to our floral units. For comparative purposes in this paper, we use the metrics of precision, recall and F1 score (see Section 3.2), and the terms F1 and detection accuracy interchangeably.

Later Xia et al. (2018) obtained a detection accuracy of 0.76 for their CNN of sea cucumber species on model-unseen internet images, and that of Tresson et al. (2019) detected and identified individuals and castes of 23 ant species to a detection accuracy of 0.88. Ditria et al. (2020) developed a CNN to estimate fish abundance that reached a detection accuracy of 0.92 in single image datasets, 7.1% better than achieved by human experts. Likewise that of Duporge et al. (2020) to estimate African elephant abundance in heterogeneous areas (weighted detection accuracy = 0.78) exceeded expert human performance by 1%.

1.4 | Faster R-CNN

Faster R-CNN (Ren et al. 2015) uses a region proposal network to rapidly find object-like regions, which are fed to a Fast R-CNN detector (Girshick, 2015) in a single model design, which has outperformed other algorithms in comparative studies on ecological data (e.g. Duporge et al., 2020; Schneider et al., 2018).

A deployed Faster R-CNN model generates region proposals by sliding a network over a test image, of tens-of-thousands of windows at differing scales and aspect ratios, assigning a binary class label (is / is not an object) at each window location. A softmax estimator of $k + 1$ classes ($k = 25$ taxa in our case, plus a 'background' class) defines a discrete probability distribution for each proposed region, and fine-tunes by regression to a final bounding box. Computation is completed in seconds, making this approach attractive for challenging fieldwork.

In this study, we use 25,000 labelled tags in almost 2000 fieldwork-realistic images to train and validate a Faster R-CNN to detect, identify and quantify the nectar-producing floral units of 25 flowering plant taxa in images of grassland in the United Kingdom. We compare the time and performance of the CNN to expert human surveyors and explore potential future applications of our approach.

2 | MATERIALS AND METHODS

2.1 | Habitat

We selected native, weed-rich grassland for automated estimation of pollinator resources due to its widespread importance in 'relaxed mowing' of urban green spaces by city councils (e.g. Scottish Government, 2019), agricultural payment-by-results schemes (Chaplin et al., 2019) and pollination research on road verges and unmown meadows (e.g. Phillips et al. 2020). Indeed Jones et al. (2021) concluded that

management changes on improved grassland have the greatest potential to increase floral resource availability across the United Kingdom.

An approximate altitudinal limit of 500 m asl was imposed to separate this habitat type from upland vegetation assemblages. We focused on nectar, although the same approach could be applied to detection of other drivers such as pollen, larval foodplants or structural features.

2.2 | Training dataset

All images were taken from vertical or near-vertical, encompassing approximately 1 m² of untrampled ground area, with no extraneous objects (e.g. quadrat frame, litter), of a maximum vegetation height of 1 m, of minimum size 2 MB and in reasonable focus. A Canon Power-shot G10 (14.7 Megapixels) was used to compile these training data.

To cover peak annual nectar production, a total of 1997 images were generated in the United Kingdom from May to August inclusive in 2019. Plants were identified using Stace (2019), though maintaining the synonym of *Jacobaea vulgaris* (*Senecio jacobaea*) for continuity with the UK Pollinator Monitoring Scheme (<https://www.growwilduk.com/blog/help-count-pollinators-science>).

We estimated that training for each taxon would require around 1000 labelled image tags, and that available time and resources allowed 25 taxa. We examined floral nectar production in this habitat (Baude et al., 2016) to refine targetting to the most common and highest nectar-producing taxa. For each taxon, we identified a floral unit definition that was practicable for image detection (Figure 1) and estimation of nectar sugar mass (Section 2.3).

2.3 | Tags and nectar values

Manual drawing and classification of bounding boxes ('tags') around floral units in the images was conducted using VoTT (Barlow, 2020), for which accurate taxonomic identification is pivotal.

To include sufficient pixels in tags, for some small-flowered taxa (e.g. *Stachys sylvatica*, *Prunella vulgaris*), our floral unit definitions comprised larger structures than were used in the original nectar assays (Baude et al., 2016; Hicks et al., 2016). For those the assay's nectar sugar mass per floral unit was multiplied accordingly.

Taxa with a similar appearance from the vertical (e.g. *Ranunculus repens* and *R. acris*; *Leontodon* spp. and *Hypochaeris radicata*) were aggregated if their nectar sugar values were not markedly divergent. The species and nectar value derivations for these aggregate taxa (*Lotus* spp., *Ranunculus* spp., yellow composites) are detailed in Table 1 and illustrated in Figure 1.

The Pascal VOC format .jpg images and .kml tag labels were checked for errors and divided into taxonomically-balanced training and validation subsets, of 1595 (20,106 tags) and 402 (5246 tags) images, respectively.



FIGURE 1 Sample tags of the 25 classes used for the CNN. In each case, the black frame in the image represents a single floral unit trained in the model. From top-left to bottom-right: *Achillea millefolium*; *Angelica sylvestris*; *Bellis perennis*; *Centaurea nigra*; *Cirsium arvense*; *C. vulgare*; *Heracleum sphondylium*; *Knautia arvensis*; *Lathyrus pratensis*; *Leucanthemum vulgare*; *Lotus* spp. (*L. corniculatus*); *Lotus* spp. (*L. pedunculatus*); *Prunella vulgaris*; *Ranunculus* spp. (*R. acris*); *Ranunculus* spp. (*R. repens*); *Rhinanthus minor*; *Rubus fruticosus* agg.; *Senecio jacobaea*; *Silene dioica/latifolia* (*S. dioica*); *Silene dioica/latifolia* (*S. latifolia*); *Stachys sylvatica*; *Symphytum officinale*; *Taraxacum* agg.; *Trifolium pratense*; *T. repens*; *Vicia cracca*; *V. sepium*; yellow composite (*Hypochaeris radicata*); yellow composite (*Leontodon hispidus*); yellow composite (*Scorzoneroides autumnalis*)

2.4 | Class balance

Flowers of some taxa (e.g. *Lathyrus pratensis*, *Taraxacum* agg.) were represented by fewer tags in the training dataset than the best represented (e.g. *Leucanthemum vulgare*, *Ranunculus* spp.), due largely to population densities in the wild. To minimize the training influence of imbalanced classes, and because we also apply image augmentations downstream (see section 2.5), we oversampled (e.g. Ferreira et al., 2019) 104 of the original images once. This increased the training dataset to 1699 images (21.934 tags. Figure S.1 in the Supporting Information).

2.5 | Model training

To train the CNN, we used the PyTorch deep learning framework (Paszke et al., 2016) based on the Torch library developed by Facebook AI Research. Training and validation files were uploaded to a Microsoft Azure Data Science Virtual Machine with the NC series CUDA-enabled GPUs available in the Western Europe region. On this virtual machine, we set up an Anaconda3 virtual environment with Python.3.8.5 (2020), PyTorch (Paszke et al., 2016), torchvision (Marcel & Rodriguez, 2010), NumPy (Harris et al., 2020), detecto (Bi, 2019) and pandas (McKinney, 2010).

To prime the Faster R-CNN, we employed transfer learning (Pan & Yang, 2010) by using available weights as initialization parameters,

in this case, a ResNet-50 Feature Pyramid Network (He et al., 2015) which was the winning architecture of the 2016 ImageNet competition (He et al., 2016). The training and prediction scripts were written in Python IDLE.3.8 (Reedy, 2020) including several manipulatable parameters, batch sizes and random augmentation of images on the training data only (horizontal flips, saturation, brightness, contrast and hue effects) to increase diversity and therefore generalization to model-unseen images. In addition, we used L2-regularization which penalizes large weights, thereby enforcing the network to use small weights to minimize model overfitting. Finally to tackle multiple overlapping proposal boxes, we applied soft non-maximum suppression (NMS; Bodla et al., 2017). This recursively sorts all detection boxes on the basis of their scores, given a user-defined proportion of intersection over union ($IoU = \text{area of overlap of boxes} / \text{area of union of boxes}$), eliminating those competing boxes with the lowest scores.

Trial models were run with a range of virtual machine specifications and PyTorch parameters, optimizing the latter to minimize the loss function on the validation set as a measure of generalization to unseen records (Figure S.2 in the Supporting Information). This is a multi-task loss function on each labelled proposal, combining the losses of classification and bounding-box regression. For those models that stabilized at a low validation loss, a range of different stopping times (= number of epochs) were trialled and the behaviour of those models on the training images was assessed visually. No lower confidence level was enforced on classifications, and an IoU value of 0.1 (10% overlap) was chosen to generalize to the majority of the taxa. The latter is aimed

TABLE 1 Species and aggregate classes in the training dataset, with derivation of nectar sugar mass per CNN floral unit (far right column). Where the floral unit definitions used in the CNN are the same as those of the original assays, nectar values are equal in the assay and model columns. Where the CNN used larger floral units the nectar values were scaled up. The CNN floral units are illustrated in Figure 1

Taxon	Original nectar assays			Conversion of nectar assays to CNN model	
	Derivation of nectar sugar mass assay value	Floral unit definition of original nectar assay	Nectar sugar mass assay value ($\mu\text{g}/\text{day}$)	Number of nectar assay floral units in each model tag	Nectar sugar mass value of each model tag ($\mu\text{g}/\text{day}$)
<i>Achillea millefolium</i> , yarrow	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	36.72	6	220.32
<i>Angelica sylvestris</i> , wild Angelica	Value of Baude et al. (2016)	Single flower	15.66	12	187.92
<i>Bellis perennis</i> , daisy	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	65.43	1	65.43
<i>Centaurea nigra</i> , black knapweed	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	6089.74	1	6089.74
<i>Cirsium arvense</i> , creeping thistle	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	4574.31	1	4574.31
<i>Cirsium vulgare</i> , spear thistle	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	8628.74	1	8628.74
<i>Heracleum sphondylium</i> , Hogweed	Value of Baude et al. (2016)	Single flower	98.17	15	1472.55
<i>Knautia arvensis</i> , field scabious	Value of Baude et al. (2016)	Inflorescence	9861.29	1	9861.29
<i>Lathyrus pratensis</i> , meadow vetchling	Value of Baude et al. (2016)	Single flower	952.69	2	1905.38
<i>Leucanthemum vulgare</i> , oxeye daisy	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	1326.36	2	2652.73
<i>Lotus</i> spp., birdsfoot trefoil aggregate	Mean value of <i>L. pedunculatus</i> and <i>L. corniculatus</i> , from Baude et al. (2016) and Hicks et al. (2016)	Single flower	53.88	4	215.52
<i>Prunella vulgaris</i> , selfheal	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Single flower	192.77	3	578.32
<i>Ranunculus</i> spp., buttercup aggregate	Mean value of <i>R. acris</i> and <i>R. repens</i> , from Baude et al. (2016) and Hicks et al. (2016)	Single flower	141.18	1	141.18
<i>Rhinanthus minor</i> , yellow rattle	Value of Baude et al. (2016)	Single flower	108.90	4	435.60
<i>Rubus fruticosus</i> , blackberry	Value of Baude et al. (2016)	Single flower	1892.83	1	1892.83
<i>Senecio jacobaea</i> , common ragwort	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	1736.44	1.5	2604.66
<i>Silene dioica/latifolia</i> , red/white campion	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Single flower	297.82	2	595.64
<i>Stachys sylvatica</i> , hedge woundwort	Value of Baude et al. (2016)	Single flower	311.11	2	622.22
<i>Symphytum officinale</i> , common comfrey	Value of Baude et al. (2016)	Single flower	1969.46	2.5	4923.65
<i>Taraxacum</i> agg., dandelion	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Capitulum	3397.82	1	3397.82
<i>Trifolium pratense</i> , red clover	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Single flower	82.61	24	1982.71
<i>Trifolium repens</i> , white clover	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Single flower	30.55	16	488.84

(Continues)

TABLE 1 (Continued)

Taxon	Original nectar assays			Conversion of nectar assays to CNN model	
	Derivation of nectar sugar mass assay value	Floral unit definition of original nectar assay	Nectar sugar mass assay value ($\mu\text{g}/\text{day}$)	Number of nectar assay floral units in each model tag	Nectar sugar mass value of each model tag ($\mu\text{g}/\text{day}$)
<i>Vicia cracca</i> , tufted vetch	Mean value of Baude et al. (2016) and Hicks et al. (2016)	Single flower	297.31	3	891.93
<i>Vicia sepium</i> , bush vetch	Value of Baude et al. (2016)	Single flower	117.07	3	351.21
Yellow composite aggregate	Mean value of <i>Hypochaeris radicata</i> , <i>Scorzoneroidea autumnalis</i> and <i>Leontodon hispidus</i> , from Baude et al. (2016) and Hicks et al. (2016)	Capitulum	1001.11	1	1001.11

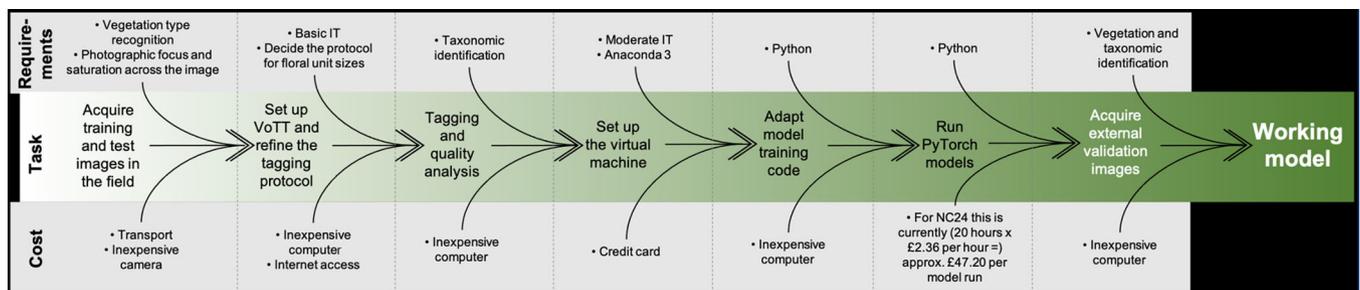


FIGURE 2 Workflow from field origin to working model

to encompass the range of floral unit presentations between crowded (e.g. Figures S.3.2 and S.3.7 in the Supporting Information) and sparse (e.g. Figures S.3.4 and S.3.12 in the Supporting Information) and is less critical to resource estimation than it would be for object enumeration.

2.6 | External validation

A test set of 50 independent images was compiled, which were as close as possible to the image specifications set out in Section 2.2. These were taken by different photographers and cameras, at different locations and countries (Table S.5 in the Supporting Information), spanning a range of 2–1346 floral units per image (Figure 3, lower panel). Note that a real survey should standardize the field of view of their quadrat images (or crop all to a marked area in the image) in order to calculate a constant nectar sugar mass per unit area. Three experienced pollination-plant surveyors were asked to count the open flowers of taxa in the images as they would in the field (minus the ability to manually interact with the plants), identify to single species or aggregates as they deemed appropriate, and to use their preferred concept of floral units (and mean nectar sugar mass values thereof) with no conferring.

2.7 | Data and analyses

Plots were produced in R.4.0.3 and RStudio.1.3.959. The raw image data, labelled tags, final object detection model and all scripts are

archived in the Dryad Digital Repository. (See Figure 2 for an indicative project workflow.)

3 | RESULTS

3.1 | Model selection

The final model (Figure S.2 in the Supporting Information: model *n*) trained using a batch size of 1699, learning rate 0.008, saturation augmentations 0.2, horizontal flip augmentations 0.4, and 50 epochs, NMS overlap (= IoU) threshold of 0.1, and took 20 h to run on a NC24 virtual machine (comprising 24 GPUs and 224 GiB RAM). The choice of a single IoU value is inevitably a compromise for 25 taxa, and for nectar sugar mass estimation we view a reduction in true positives as worse than an increase in false positives. The final model specifications were run twice to confirm stability of the algorithm's output with respect to its stochastic training.

The model outputs visual displays of its proposal regions and taxon classifications (Figure S.3 in the Supporting Information), and a spreadsheet of counts of detected taxa converted to nectar sugar mass using the floral unit concept of the model (Table 1). On a standard outdated laptop (i5-4200M CPU, 4 GiB RAM), this takes approximately 5.5 s.

3.2 | Performance

For performance assessment we define (i) true positives as those boxes with correct detection and classification of floral units by the model; (ii) false positives as those model-detected floral units where there are either none in the image or where a floral unit of a different taxon is present; (iii) false negatives as floral units in the image which are not detected by the model; and (iv) true negatives as an incalculable high number. A confusion matrix was constructed from the 50 test images (Figure S.4 in the Supporting Information) which illustrates performance per taxon on this dataset. Key performance metrics for the final model are as follows:

$$\begin{aligned} \text{Precision of floral units} &= \text{True Positive} / (\text{True Positive} + \text{False Positive}) \\ &= 926 / (926 + 101) \\ &= 0.90 \end{aligned}$$

$$\begin{aligned} \text{Recall of floral units} &= \text{True Positive} / (\text{True Positive} + \text{False Negative}) \\ &= 926 / (926 + 152) \\ &= 0.86 \end{aligned}$$

$$\begin{aligned} \text{Mean F1 score} &= 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \\ &= 2 \times (0.90166 \times 0.85900) / (0.90166 + 0.85900) \\ &= 0.88 \end{aligned}$$

These scores indicate that the model makes floral unit detection errors at an average rate of 1–2 in 10. Single taxon F1 scores in this limited test set range from 0.61 (*Achillea millefolium*) to 0.95 (*Leucanthemum vulgare*).

3.3 | Comparison to human surveyors: Single quadrats

The CNN estimates of nectar sugar mass on the 50 model-unseen images are plotted against those of the three human surveyors in Figure 3 (upper panel), wherein 27 of the 50 CNN estimates are within the absolute range of the human estimates. The OLS regression line has intercept near zero, gradient near identity, and its proportion of variance explained approaches 95% (see Figure 3 legend for parameter values and Table S.5 for raw data in the Supporting Information). Mean estimation time per image was 5.5 s by CNN and 3.2 min by human.

The same exercise for total floral units returned an inferior fit (Figure 3 lower panel). The average within-image standard deviation for these three surveyors for total nectar sugar mass was 5.43 mg/day and for total floral units was 34.

3.4 | Comparison to human surveyors: Replicated quadrats

For sampling vegetation, surveyors generally take the mean value of an optimized number of quadrats per site. One way to identify an optimal protocol is to determine the number of quadrats (in our case, images) to achieve a stable estimate for the parameter of interest (in our case, nectar sugar mass per image). For this weed-rich grassland vegetation, we generated 30 randomized orderings (without replacement) of the whole set of 50 images, in each case generating the cumulative mean nectar sugar mass for each additional increments of one image. While our test set was not a formal ecological pilot, the nectar sugar mass estimates are of similar field of view (1–3 m²) in each case. The expectation is that estimation variability across random sequences demonstrates a sample size above which estimates remain stable.

On the human nectar sugar mass estimates (Figure 4, upper panel), there appears to be a cost–benefit ‘elbow’ at around 15 and 20 samples, where the mean is converging on approximately 27 mg per day, the absolute range of variability of the randomized estimates at approximately 35, and the width of two standard deviations between 10 and 20.

A similar pattern, mean and variability on the y-axis are apparent for the CNN nectar sugar mass estimates (Figure 4, lower panel). Both estimators thus indicate the same sampling protocol.

As the total possible sample combinations of this method (given by $C(n,r) = n! / (r!(n-r)!)$, where $n = 50$ and $r = \text{image number}$) at 15 and 20 samples are 2.3×10^{12} and 4.7×10^{13} , respectively, a stable standard deviation is considered to be robust to field reality.

4 | DISCUSSION

We constructed a CNN model to assess daily nectar sugar mass per quadrat image in UK native weed-rich grassland across a wide range of floral abundance, which identifies and counts floral units in this habitat 35 times more rapidly than an expert human surveyor.

The model makes floral unit detection errors at a rate of 1–2 in 10, and we would not recommend this method for identification. For total nectar sugar mass of single images, most of the model's estimates fall within the absolute range of those obtained by human surveyors. The full benefit of the model becomes apparent when estimating vegetation properties from multiple quadrat images, for which the model's mean and variation closely match those of human surveyors. Furthermore, the estimate of the model is deterministic and therefore has known subjectivity, contrasting to intra- and inter-human survey variability.

This has particular applicability to the grading of agri-environment schemes, as these move from conventional management prescriptions to a scaled link between payment and results (Chaplin et al., 2019) in Europe and the United Kingdom. We focused on grassland due to this habitat's high-scoring natural capital (Dicks et al., 2015) and potential

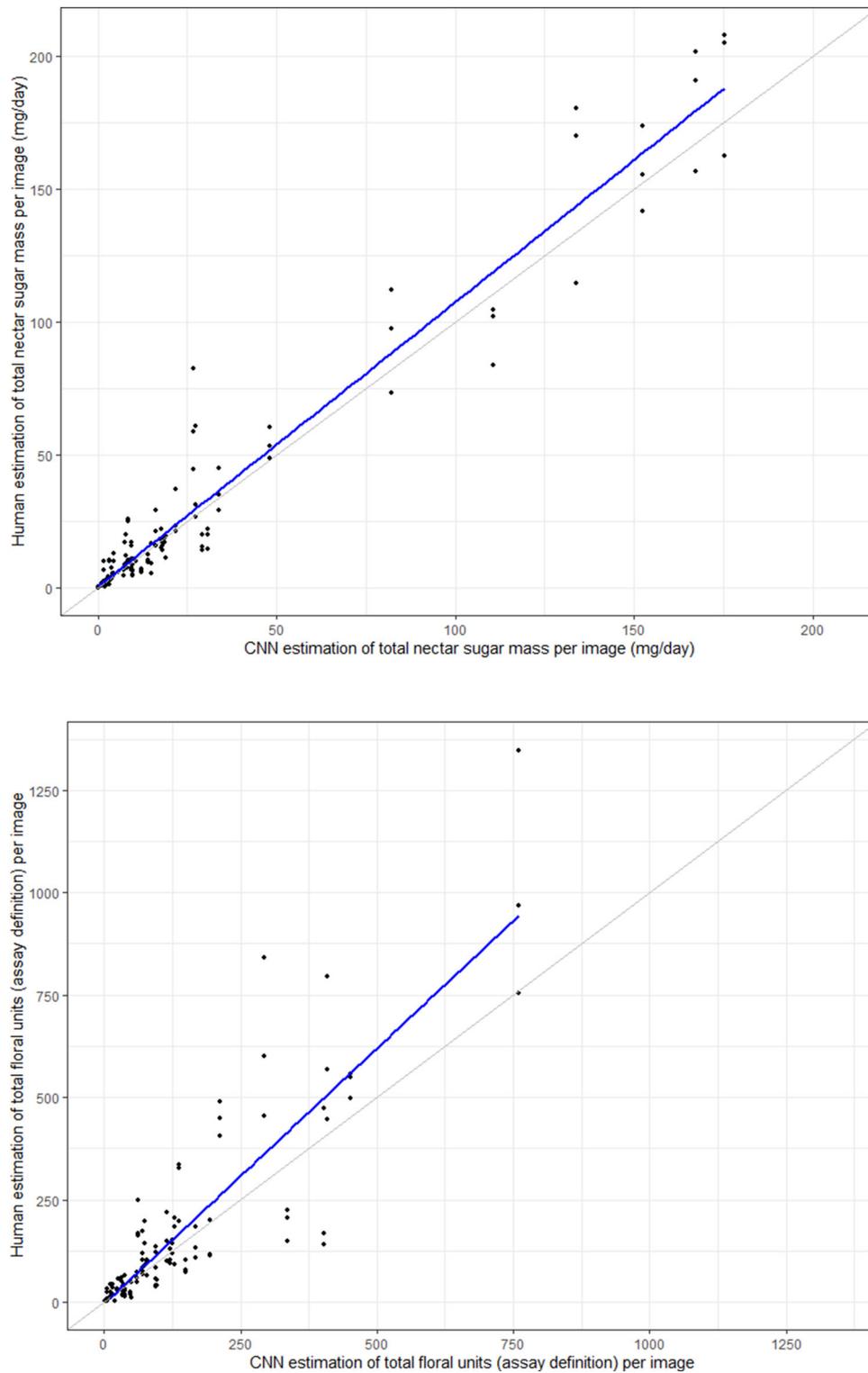


FIGURE 3 External validation of CNN against a test set of 50 images, of total nectar sugar mass per day (upper panel: $y = 1.07x + 0.43$; $r^2 = 0.95$) and total floral units (lower panel: $y = 1.25x + 3.66$; $r^2 = 0.78$) by three independent pollinator-plant surveyors. OLS lines marked blue. Identity $x = y$ lines marked grey

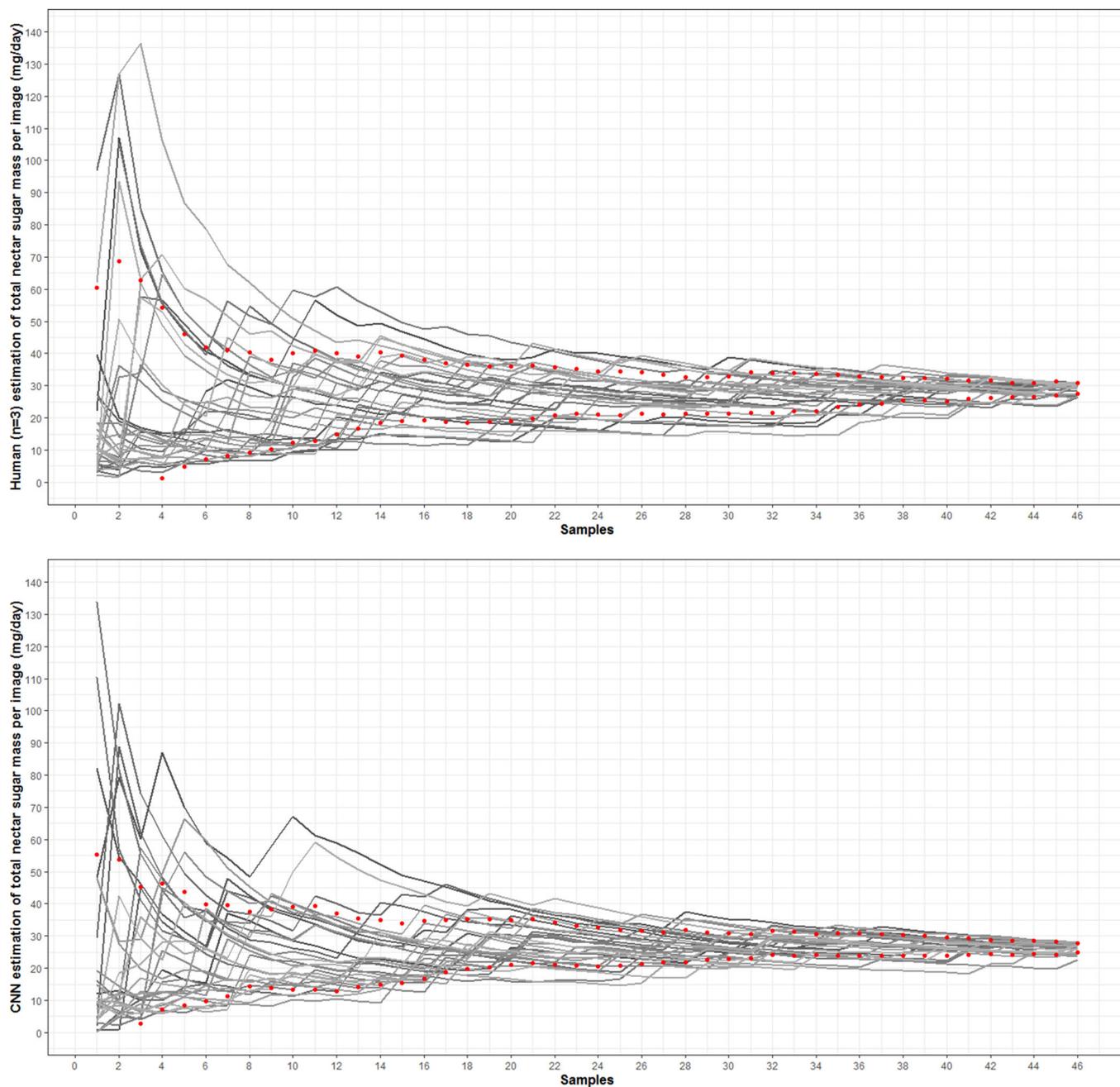


FIGURE 4 Thirty randomization runs of mean ($n = 3$) human-estimated (upper panel) and CNN-estimated (lower panel) total nectar sugar mass per image to inform sample number, up to a 46-sample protocol. Red dots mark the limits of one standard deviation either side of the sample means

for improvement of its ecosystem service (Baude et al., 2016), but would envisage the same strategy for any terrestrial or aquatic habitat that can be photographed.

Though the time taken to generate training data for the model is considerable (Table S.6 in the Supporting Information), this method could cut floral survey time per stand from hours to minutes. The same could be adapted to other image-based vegetation resources, such as winter bird food, pollen resources, insect infestation and tree flower/fruit abundance, and would require minimal additional coding to adapt to live video input from a phone, Raspberry Pi or

drone, in order to survey the entire stand rather than a sample of quadrats.

4.1 | Object detection of floral units

Object detection is imperfect, and there are several drivers of errors, underlining the importance of exposing the model to external data. Xia et al. (2018) found that discrimination of sea cucumber species declined from 0.98 for their training set to 0.76 for model-unseen

internet images; Xu and Matzner (2018) concluded that their local fish species model could not reliably recognize fish in real-world applications.

NMS can reject valid object proposals if set too high – a more likely scenario when floral units are densely packed in an image – or falsely include overlapping object proposals if set too low. For nectar sugar mass estimation, we took the view that a reduction in true positives is a worse outcome than an increase in false positives, particularly given that NMS rejects low-confidence proposals which are not singletons in uninterrupted horizontal space. We set no lower confidence limit on the model's classifications, though it would be straightforward to do so by thresholding the output, and for surveyors who are solely interested in presence-absence of taxa this may be a profitable route to explore. Such an approach could experimentally determine an optimal confidence threshold by finding that which maximizes the F1 score, in the manner of Kawazoe et al. (2018).

4.2 | Taxonomic and structural definitions

Aggregation of taxa (e.g. *Ranunculus* spp., *Lotus* spp.; Table 1) suffices when the nectar values for the species in each aggregate taxon do not differ widely. This was designed in to our image tagging, yet there remain cases of species which are morphologically similar from a vertical view but produce significantly divergent masses of nectar sugar. Of our 25 taxa the most difficult to differentiate are yellow Asteraceae. Of these, dandelions (*Taraxacum* agg.) provide 3.4 times as much nectar sugar as our species aggregate 'yellow composite', yet *Taraxacum* agg. can be incorrectly classified as yellow composite (Figure S.3.4 in the Supporting Information). In our test set, this error occurred at a rate of 50%, while the converse error (yellow composite incorrectly classified as *Taraxacum* agg.) was not encountered. This is a general problem when taxa are phenotypically similar, and its solution could be challenging for an object detection model of quadrat images. In another pollination-related study, the CNN of Hansen et al. (2020) on flower-visiting insects experienced a drop in recall (the proportion of actual positives identified correctly) from 0.75 at aggregate (genus) level to 0.53 at species level, even with museum specimen images.

For some species, there is the opposite problem, of failing to correctly identify variable instances of it as a single taxon. Species such as *S. dioica*, *Stachys sylvatica* and *Vicia* spp. have variably oriented pedicels that (in contrast to vertically oriented flowers such as the Asteraceae considered above) present a heterogeneous view from different aspects (e.g. Figures S.3.8 and S.3.9 in the Supporting Information). For such taxa, a comprehensive training set is likely to be critical. We considered the deep learning approach to be the most powerful available for determining common features from such highly variable training class images.

Floral unit definition matters for nectar sugar mass estimation, and considerable intraspecific variation in floral resource production necessitates average values (Dicks et al., 2015) for the pollinator-visited plants in any given survey. Experienced surveyors aggregate spatial patterns differently (Figure 3, lower panel), and a standardiza-

tion of floral units such as that imposed by a CNN may guard against them being variably defined by different studies.

For a single quadrat, resource estimation would be simpler if open flower number per floral unit differed less among plants, soils and seasons (e.g. *Trifolium* spp. (Figures S.3.1 and S.3.12 in the Supporting Information) and *Heracleum sphondylium* (Figure S.3.3. in the Supporting Information)). This variation is captured somewhat by using mean nectar sugar mass values from multiple individuals sampled (in 16 of our taxa; Table 1) in at least two studies and populations. These challenges are general when allocating average reward estimates to flowers, not restricted to automated detection, but it is likely that a CNN will tend to require larger floral units and therefore more assumptions, at least with field images of a practical resolution.

By necessity, the floral units of diminutive plants (e.g. *Bellis perennis*, *Trifolium repens*, *Prunella vulgaris*) have small bounding boxes comprising less pixel information, so for these taxa the surveyor's photographic focus is particularly important. Low image quality will confound both object detection and identification, just as it does for humans. While our image specifications are aimed to be strict, floral units may be obscured in vertically stratified vegetation, and achieving crisp focus and zero washout for all flowers in the three-dimensional space of a quadrat may be challenging.

4.3 | Target taxa

The obvious practical limitation to an object detection model is that it is trained on a limited set of taxa, and that the inclusion of an object class 'other' (meaning detected floral units not allocated to model-defined taxa) is in many habitats intractable. For the intended purpose of nectar sugar estimation therefore, a surveyor would want to be reasonably confident that their habitat does not include non-target taxa in abundance at the scale of square metres. For model training, we aimed to include all high nectar-producing taxa that appear in this habitat and have not tested how well the model would perform where floral units of non-target taxa are present. As those floral units would likely be detected and classified as one that the model knows, their nectar contribution to quadrat estimates would be based on incorrect species values. If such non-target species are rare in a quadrat, and/or produce nectar per floral unit similar to target taxa, the impact of these errors is likely to be small and quadrat level estimates of nectar sugar mass would be little affected. Given the time taken on the image-labelling step, we estimate that every taxon adds around three person-weeks to project time (Table S.6 in the Supporting Information), although there are promising signs from other work (e.g. Arazo et al., 2019) that an unsupervised model could be used at labelling stage so to accelerate this.

Detection performance varies by taxon, in part due to the requirement to optimize a single confidence threshold for 25 taxa, whose test recall values range from 0.50 (*Achillea millefolium*) to 0.95 (*Leucanthemum vulgare*). This variation parallels that of Hansen et al. (2020), whose CNN reported recall values of between 0.20 and 1 for 80 beetle genera. Ideally, we would want to set a different threshold per taxon,

being higher for those morphologically-variable floral units (e.g. *Lathyrus pratensis*, *Silene dioica*) for which the model's concept can be over-inclusive, and lower for those morphologically uniform floral units (e.g. *Ranunculus* spp., *Trifolium pratense*) which tend to be more precisely detected. We do not advise using the model as an object detection tool alone for the whole set of 25 taxa, but for a subset of these (e.g. *Leucanthemum vulgare*, *Trifolium repens*, *Trifolium pratense*, yellow composite) or a low diversity habitat (e.g. road verges, city parks) its performance may be sufficient. For such a smaller set of classes, the user could adjust both the confidence threshold and NMS value to those target taxa.

4.4 | Further directions

Some additional modelling exercises might achieve small gains in object detection performance. Ensembling of models has been successful elsewhere (e.g. Nobashi et al., 2020; Priyadarshini & Puri, 2021), and actually would be fairer in our model tests which effectively pitted a single CNN against an ensemble of human surveyors. Some experimentation with different modelling architectures and other CNN algorithms may reveal one that performs better in this case, and training two models simultaneously on this dataset as generative adversarial networks (Goodfellow et al., 2014) would be an informative test of our choice of deep learning model.

ACKNOWLEDGEMENTS

Many thanks to Chris Monk, David Whiting, all at Microsoft AI-for-Earth, Alan Bi, Lucy Hatcher, Andy Cooke, Vicky Robinson, Annabelle LePage, Sarah Escott, Babs Pearson, Abbie Keeper, Colin Prosser, Richard Alexander, Howard van Rooijen, Daveron Smith, Rebecca Jackson-Pitt and Mary Vagiou. This work was funded by Natural England and Microsoft AI-for-Earth, with a set of additional images contributed by the POLLEN project (2016-00108486) of the Centre-Val de Loire.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS' CONTRIBUTIONS

DH, CK, GS and MB designed the project; MB, PO and DH compiled the image datasets; DH and CK analysed the data; All authors contributed critically to the manuscript drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

The original images (.jpg), labelled tags (PASCAL VOC format) and model (.th Torchvision format) are available from the Dryad Digital Repository <https://doi.org/10.5061/dryad.63xsj3v34> (Hicks et al., 2021a). The Python and R codes are available in Zenodo <https://doi.org/10.5281/zenodo.5347794> (Hicks et al., 2021b).

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1002/2688-8319.12099>

ORCID

Damien Hicks  <https://orcid.org/0000-0001-7884-5239>

Pierre Ouvrard  <https://orcid.org/0000-0002-2805-4583>

REFERENCES

- Arazo, E., O'Connor, N. & McGuinness, K. (2019). Improving unsupervised learning with exemplar CNNs. In *IMVIP 2019: Irish Machine Vision & Image Processing*. 28–30 August 2019. Technological University Dublin. <https://doi.org/10.21427/9mp9-0t26>
- Baldock, K. C. R., Goddard, M. A., Hicks, D. M., Kunin, W. E., Mitschunas, N., Morse, H., Osgathorpe, L. M., Potts, S. G., Robertson, K. M., Scott, A. V., Staniczenko, P. P. A., Stone, G. M., Vaughan, I. P. & Memmott, J. (2019). A systems approach reveals urban pollinator hotspots and conservation opportunities. *Nature Ecology and Evolution* 3, 363–373. <https://www.nature.com/articles/s41559-018-0769-y>
- Baldock, K. C. R., Goddard, M. A., Hicks, D. M., Kunin, W. E., Mitschunas, N., Osgathorpe, L. M., Potts, S. G., Robertson, K. M., Scott, A. V., Stone, G. M., Vaughan, I. P. & Memmott, J. (2015). Where is the UK's pollinator biodiversity? The importance of urban areas for flower-visiting insects. *Proceedings of the Royal Society B*. 282, 1803. <http://rspb.royalsocietypublishing.org/content/282/1803/20142849>
- Barlow, T. (2020). Visual object tagging tool. Commercial Software Engineering (CSE) group, Microsoft. <https://github.com/microsoft/VoTT>
- Baude, M., Kunin W. E., Boatman, N. D., Conyers, S., Davies, N., Gillespie, M. A. K., Morton, D., Smart, S. M., & Memmott, J. (2016). Historical nectar assessment reveals the fall and rise of floral resources in Britain. *Nature* 530, 85–88. <https://www.nature.com/articles/nature16532>
- Bengio, Y. (2012). Deep learning of representations for unsupervised and transfer learning. *JMLR Workshop and Conference Proceedings* 27, 17–37. <http://proceedings.mlr.press/v27/bengio12a/bengio12a.pdf>
- Bi, A. (2019). Detecto for PyTorch. <https://detecto.readthedocs.io/en/latest/usage/quickstart.html#technical-requirements>
- Bodla, N., Singh, B., Chellappa, R. & Davis, L. (2017). Improving object detection with one line of code. *arXiv*, 1704.04503v2
- Breeze T. D., Bailey, A. P., Balcombe, K. G., Brereton, T., Comont, R. Edwards, M., Garratt, M. P., Harvey, M. Hawes, C., Isaac, N., Jitlal, M., Jones, C. M., Kunin, W. E., Lee, P., Morris, R. K. A., Musgrove, A., O'Connor, R. S., Peyton, J., Potts, S. G., Roberts, S. P. M., Roy, D. B., Roy, H. E., Tang, C. Q., Vanbergen, A. J. & Carvell, C. (2020). Pollinator monitoring more than pays for itself. *Journal of Applied Ecology* 58(1), 44–57. <https://doi.org/10.1111/1365-2664.13755>
- Carvell, C., Meek, W. R., Pywell, R. F. & Nowakowski, M. (2004). The response of bumblebees to successional change in newly created arable field margins. *Biological Conservation*, 118, 327–339
- Chaplin, S., Robinson, V., LePage, A., Keep, H., Le Cocq, J., Ward, D., Hicks, D., & Scholz, E., (2019). *Pilot results-based payment approaches for agri-environment schemes in arable and upland grassland systems in England. Final Report to the European Commission*. Natural England and Yorkshire Dales National Park Authority. <http://publications.naturalengland.org.uk/publication/6331879051755520>
- Chen, G., Han, T. X., He, Z., Kays, R. & Forrester, T. (2014). Deep convolutional neural network based species recognition for wild animal monitoring. Paper presented at the 2014 IEEE International Conference on Image Processing (ICIP), IEEE, New York.
- Dicks, L. V., Baude, M., Roberts, S. P. M., Phillips, J., Green, M. & Carvell, C. (2015). How much flower-rich habitat is enough for wild pollinators? Answering a key policy question with incomplete knowledge. *Ecological Entomology* 40(1), 22–35. <https://doi.org/10.1111/een.12226>
- Ditria, E., Lopez-Marcano, S., Sievers, M., Jinks, E., Brown, C. & Connolly, R. (2020). Automating the analysis of fish abundance using object detection: Optimizing animal ecology with deep learning. *Frontiers in Marine Science* 7, 429. <https://doi.org/10.3389/fmars.2020.00429>
- Duporge, I., Isupova, O., Reece, S., Macdonald, D. & Wang, T. (2020). Using very-high-resolution satellite imagery and deep learning to detect and

- count African elephants in heterogeneous landscapes. *Remote Sensing in Ecology and Conservation*. Advance online publication. <https://doi.org/10.1002/rse2.195>
- Ferreira, A., Silva, L., Renna, F., Brandl, H., Renoult, J., Farine, D., Covas, R. & Doutrelant, C. (2019). Deep learning-based methods for individual recognition in small birds. *Methods in Ecology and Evolution*, 11, 1072–1085. <https://doi.org/10.1111/2041-210X.13436>
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation, Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Girshick, R. (2015). Fast R-CNN. Paper presented at IEEE International Conference on Computer Vision (ICCV). arXiv:1504.08083v2
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Bengio, Y. (2014). Generative adversarial networks. *Advances in Neural Information Processing Systems* 3, 2672–2680. <https://arxiv.org/abs/1406.2661>
- Hansen O. L. P., Svenning J. - C., Olsen K., Dupont, S., Garner, B. H., Iosifidis, A., Price, B. & Høye, T. T. (2020). Species-level image classification with convolutional neural network enables insect identification from habitus images. *Ecology and Evolution* 10, 737–747. <https://doi.org/10.1002/ece3.5921>
- Harris, C. R., Millman, K. J., van der Walt, S. J. (2020). Array programming with NumPy. *Nature* 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>.
- He, K., Zhang, X., Ren, S. & Sun, J. (2015). Deep residual learning for image recognition. *Microsoft Research*. ArXiv:1512.03385.
- He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep residual learning for image recognition. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Hicks, D. M., Baude, M., Kratz, C., Ouvrard, P. & Stone, G. N. (2021a). Data from: Deep learning object detection to estimate the nectar sugar mass of flowering vegetation. *Ecological Solutions and Evidence*. Dryad Digital Repository. <https://doi.org/10.5061/dryad.63xsj3v34>
- Hicks, D. M., Baude, M., Kratz, C., Ouvrard, P. & Stone, G. N. (2021b). Deep learning object detection to estimate the nectar sugar mass of flowering vegetation. Zenodo. <https://doi.org/10.5281/zenodo.5347794>
- Hicks, D. M., Ouvrard, P., Baldock K. C. R., Baude, M., Goddard, M. A., Kunin, W. E., Mitschunas, N., Memmot, H., Morse, H., Nikolitsi, M., Osgathorpe, L. M., Potts, S. G., Robertson, K. M., Scott, A. V., Sinclair, F., Westbury, D. B. & Stone, G. N. (2016). Food for pollinators: quantifying the nectar and pollen resources of urban flower meadows. *PLoS ONE* 11(6), e0158117. <https://doi.org/10.1371/journal.pone.0158117>
- Jones, L., Brennan, G. L., Lowe, A., Creer, S., Ford, C. R. & de Vere, N. (2021). Shifts in honeybee foraging reveal historical changes in floral resources. *Nature Communications Biology* 4, 37. <https://www.nature.com/articles/s42003-020-01562-4>
- Kawazoe, Y., Shimamoto, K., Yamaguchi, R., Shintani-Domoto, Y., Uozaki, H., Fukayama, M. & Ohe, K. (2018). Faster R-CNN-based glomerular detection in multistained human whole slide images. *Journal of Imaging* 4, 91. <https://www.mdpi.com/2313-433X/4/7/91>
- LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature* 521(7553), 436–444.
- Marcel, S. & Rodriguez, Y. (2010). Torchvision the machine-vision package of torch. In Proceedings of the 18th ACM International Conference on Multimedia (pp. 1485–1488). ACM. <https://doi.org/10.1145/1873951.1874254>
- McKinney, W. (2010). pandas. In Proceedings of the 9th Python in Science Conference (p. 445). Zenodo. <https://doi.org/10.5281/zenodo.3509134>
- Nobashi, T., Zacharias, C., Ellis, J. K., Ferri, V., Koran, M. E., Franc, B. L., Iagaru, A., & Davidzon, G. A. (2020). Performance comparison of individual and ensemble CNN models for the classification of brain 18F-FDG-PET scans. *Journal of Digital Imaging* 33, 447–455. <https://doi.org/10.1007/s10278-019-00289-x>
- Norouzzadeh, S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *PNAS* 115, 25. www.pnas.org/cgi/doi/10.1073/pnas.1719367115
- Ouvrard, P., Transon, J., & Jacquemart, A. L. (2018). Flower-strip agri-environment schemes provide diverse and valuable summer flower resources for pollinating insects. *Biodiversity and Conservation* 27, 2193–2216. <https://doi.org/10.1007/s10531-018-1531-0>
- Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
- Paszke, A., Gross, S., Chintala, S., & Chanan, G. (2016). PyTorch. Facebook AI Research (FAIR). <https://pytorch.org/>
- Phillips, B. B., Wallace, C., Roberts, B. R., Whitehouse, A. T., Gaston, K. J., Bullock, J. M., Dicks, L. V., & Osborne, J. L. (2020). Enhancing road verges to aid pollinator conservation: A review. *Biological Conservation*, 250, 108687. <https://doi.org/10.1016/j.biocon.2020.108687>
- Priyadarshini, I., & Puri, V. (2021). A convolutional neural network (CNN) based ensemble model for exoplanet detection. *Earth Science Informatics*, 14, 735–747. <https://doi.org/10.1007/s12145-021-00579-5>
- Python Software Foundation. (2020). Python 3.8.5. <https://www.python.org/downloads/release/python-385/>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- RStudio Team. (2020). *RStudio: Integrated development environment for R*. RStudio. <http://www.rstudio.com/>
- Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (pp. 779–788). IEEE Press.
- Reedy, T. J. (2020). *Python integrated development and learning environment*. <https://github.com/python/cpython/tree/3.9/Lib/idlelib/>
- Ren, S., He, K., Girshick, R. & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks, Paper presented at the Neural Information Processing Systems (NIPS, 2015). (<https://arxiv.org/abs/1506.01497>)
- Schneider, S., Tayloer, G. & Kremer, S. (2018). Deep learning object detection methods for ecological camera trap data. Paper presented at the Conference of Computer and Robot Vision, 2018, Cornell University. arXiv:1803.10842
- Scottish Government. (2019). *Pollinator strategy for Scotland: 2019 progress report*. https://www.nature.scot/sites/default/files/2020-02/Pollinator%20Strategy%20for%20Scotland%20-%202019%20Progress%20Report_0.pdf
- Stace, C. A. (2019). *New flora of the British Isles* (4th ed). C & M Floristics.
- Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., Sweeney, S. J., Vercauteren, K. C., Snow, N. P., Halseth, J. M., Di Salvo, P. A., Lewis, J. S., White, M. D., Teton, B., Beasley, J. C., Schlichting, P. E., Boughton, R. K., Wight, B., Newkirk, E. S., Ivan, J. S., Odell, E. A., Brook, R. K., Lukacs, P. M., Moeller, A. K., Mandeville, E. G., Clune, J., & Miller, R. S. (2019). Machine learning to classify animal species in camera trap images: Applications in ecology. *Methods in Ecology and Evolution* 10, 585–590. <https://doi.org/10.1111/2041-210X.13120>
- Torney, C. J., Lloyd-Jones, D. J., Chevallier, M., Moyer, D. C., Maliti, H. T., & Mwitia, M. (2019). A comparison of deep learning and citizen science techniques for counting wildlife in aerial survey images. *Methods in Ecology and Evolution* 10, 779–787. <https://doi.org/10.1111/2041-210X.13165>
- Tresson, P., Tixier, P., Puech, W. & Carval, D. (2019). Insect interaction analysis based on object detection and CNN. In 2019 IEEE 21st International Workshop on Multimedia Signal Processing (IEEE MMSp), Kuala Lumpur, Malaysia (pp. 1–6). IEEE. hal-02361210
- Timberlake, T. P., Vaughan, I. P., Baude, M. & Memmott, J. (2021). Bumblebee colony density on farmland is influenced by late-summer nectar supply

- and garden cover. *Journal of Applied Ecology* 58(5), 1006–1016. <https://doi.org/10.1111/1365-2664.13826>
- Valan, M., Makonyi, K., Maki, A., Vondráček, D. & Ronquist, F. (2019). Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. *Systematic Biology* 68(6), 876–895. <https://doi.org/10.1093/sysbio/syz014>
- Villon, S., Mouillot, D., Chaumont, M., Darling, E. S., Subsol, G. & Claverie, T. (2018). A deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecological Informatics* 48, 238–244. <https://doi.org/10.1016/j.ecoinf.2018.09.007>
- Xia, C., Fu, L., Liu, H. & Chen, L. (2018). In situ sea cucumber detection based on deep learning approach. In Proceedings of the 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (pp. 1–4). IEEE.
- Xu, W., Matzner, S. (2018). *Underwater fish detection using deep learning for water power applications*. Cornell University arXivLabs. <https://arxiv.org/abs/1811.01494>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: 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>