ARTICLES
https://doi.org/10.1038/s41893-019-0246-x
nature sustainability

Deep learning to map concentrated animal feeding operations

Cassandra Handan-Nader[1,2] and Daniel E. Ho[1,2,3]*

Enforcement of environmental law depends critically on permitting and monitoring intensive animal agricultural facilities, known in the United States as 'concentrated animal feeding operations' (CAFOs). The current legal landscape in the United States has made it difficult for government agencies, environmental groups and the public to know where such facilities are located. Numerous groups have, as a result, conducted manual, resource-intensive enumerations based on maps or ground investigation to identify facilities. Here we show that applying a deep convolutional neural network to high-resolution satellite images offers an effective, highly accurate and lower cost approach to detecting CAFO locations. In North Carolina, the algorithm is able to detect 589 additional poultry CAFOs, representing an increase of 15% from the baseline that was detected through manual enumeration. We show how the approach scales over geography and time, and can inform compliance and monitoring priorities.

As US agriculture has grown and industrialized, livestock production has shifted to operations that exclusively raise large numbers of animals in confinement[1]. These CAFOs are estimated to produce more than 40% of US livestock[2], with large facilities each raising over 2,500 pigs or 125,000 broiler chickens[3]. Such specialization not only increases productivity, but also raises important questions about environmental impact. For recent reviews on the evidence of health and environmental effects of CAFOs, see previously published studies[4,5]. As the Environmental Protection Agency (EPA) noted, agriculture is "the leading contributor of pollutants to identified water quality impairments in the Nation's rivers and streams, lakes, ponds, and reservoirs"[6]. Livestock generate 13 to 25 times the amount of manure that humans produce; however, animal waste, in contrast to human waste, is not required to be treated[7]. CAFOs generate about 335 million tons of waste per year, with excess nutrients posing considerable ecological and human health risks[8,9].

Yet litigation over the past few decades has hindered regulatory efforts and environmental monitoring of CAFOs. Under the US Clean Water Act, CAFOs that 'discharge' pollutants into national waters are required to apply for and abide by the terms of a federal permit, known as the National Pollutant Discharge Elimination System (NPDES). NPDES permits require discharge reporting and violations of permit effluent limitations can trigger substantial fines. Federal courts have rejected the EPA's proposals to cover CAFOs that had the potential to discharge, holding that the EPA could only require permits for CAFOs that actually discharge pollutants (see *Waterkeeper v. EPA*[10] and *Nat'l Pork Producers Council v. EPA*[11]). Under-permitting—the failure to apply for a permit when required—is known to be a serious issue, as the EPA estimates that nearly 60% of CAFOs do not hold permits[12]. In 2008, the Government Accountability Office lamented the lack of basic information about CAFOs: "no federal agency collects accurate and consistent data on the number, size, and location of CAFOs"[13]. In 2011, the EPA proposed running a survey to collect such information, including global positioning system (GPS) location and size. The EPA sought comments on existing data sources, including the possibility of "augment[ing] information from satellite images and aerial photography location information to obtain a comprehensive, consistent national inventory of CAFOs"[12]. The proposal was withdrawn in 2012, with the EPA highlighting the possibility of obtaining such information from existing data sources[14]. In 2016, after the EPA requested the return of data it had released, a federal court concluded that the EPA could not release available GPS coordinates as a dataset, even in instances where individual locations were publicly available (*American Farm Bureau Fed. v. EPA*[15]). Owing to these developments, scholars have criticized the lack of basic information about CAFOs[16–18].

In response to the lack of information about CAFOs, a wide range of environmental and public interest groups—including Pew Charitable Trusts[19,20], the National Resources Defense Council[20], the Environmental Working Group[21], Food and Water Watch[22], Waterkeeper Alliance[21], the Sierra Club[23] and EarthJustice[20,24]—has engaged in extensive and resource-intensive efforts to improve monitoring of CAFOs. Because location information is critical for monitoring discharges and other forms of environmental noncompliance, numerous groups have attempted to collect more-comprehensive information about CAFO coordinates. Two groups, for instance, hired contractors to manually scan through satellite images in select states to identify CAFOs[21], including over three years of data for each state (S. Rundquist, personal communication). In other instances, contractors physically identified facilities by plane or car (J. Quinlivan, personal communication). After environmental interest groups (EIGs) petitioned to remove Iowa's authority to permit CAFOs, the EPA and Iowa's Department of Natural Resources entered into an agreement requiring Iowa to conduct a comprehensive census of CAFOs. To comply with the agreement, Iowa employed individuals to manually label satellite images, a process that took roughly 3.5 years to complete[25]. Although valuable, such efforts are resource-intensive and do not easily scale over time and geography.

We contribute to this important issue of public health and environmental law, science and sustainability by demonstrating how recent advances in image-learning, or 'deep-learning', techniques

[1]Stanford Law School, Stanford University, Stanford, CA, USA. [2]Department of Political Science, Stanford University, Stanford, CA, USA. [3]Stanford Institute for Economic Policy Research, Stanford, CA, USA. *e-mail: dho@law.stanford.edu

Fig. 1 | Examples of image occlusion. a–d, Examples of occluded (**a**,**c**) and non-occluded (**b**,**d**) images of poultry (**a**,**b**) and pig (**c**,**d**) CAFOs.

can considerably lower the cost of environmental monitoring and regulatory enforcement. Applications of deep learning to agricultural data have been reviewed previously[26]; however, the previous study does not include a case for environmental regulation of CAFOs. Although locating facilities alone does not measure discharges, it is a necessary, and time-consuming, step for any form of monitoring. We construct a dataset of pig and poultry CAFOs in North Carolina, by manually validating the data from a manual census conducted by two leading EIGs against publicly available, high-resolution satellite images from the US Department of Agriculture's National Agricultural Imagery Program (NAIP). Acquired annually during the growing season on a three-year cycle staggered across states, these images are easily downloadable through an online service at resolutions up to 1 m per pixel (Supplementary Notes). We downloaded all imagery for the state of North Carolina in the format of 299 × 299 image tiles at a resolution of 1 m per pixel and manually tagged 24,440 images. We trained two convolutional neural networks (CNNs) to detect the presence of pig and poultry CAFOs individually and demonstrate high-classification accuracy on the 25% of image tiles reserved for testing.

Although a necessary first step, determining whether an image tile contains a CAFO (henceforth referred to as 'image-level' predictions) is less useful for the regulatory context, as a single CAFO facility can be split across multiple image tiles. To consolidate image scores into actual facility locations for regulators (henceforth referred to as 'facility-level' predictions), we focused on poultry operations, which can contribute as much nutrient run-off to watersheds as pig operations[27,28], but are largely unpermitted in North Carolina and are therefore substantially harder to detect. We developed an algorithm to transform scores for image tiles into latitude and longitude point locations for individual poultry facilities. We show with manually validated results that we can detect 15% more poultry CAFOs than previously known. We further illustrate the utility of our approach by showing how it scales geographically and longitudinally and provides legally important estimates of facility size and other compliance priorities. Given the global growth in intensive livestock farming[29], our methods have potential for other jurisdictions.

Results

First, we demonstrate the accuracy of the model at the image level on the 25% of manually validated images that was reserved for test-

Table 1 | Image accuracy statistics

	Area under the curve			CAFO	Control
	Image	ROC	PR		
Poultry	All images	0.972	0.917	833	4,871
	Non-occluded	0.998	0.980	262	4,871
Pigs	All images	0.986	0.923	399	4,871
	Non-occluded	1.000	0.991	124	4,871

Statistics are presented for all test images (all images) and test images without occluded CAFO images (non-occluded). PR, precision-recall curve; ROC, receiver-operating characteristic curve.

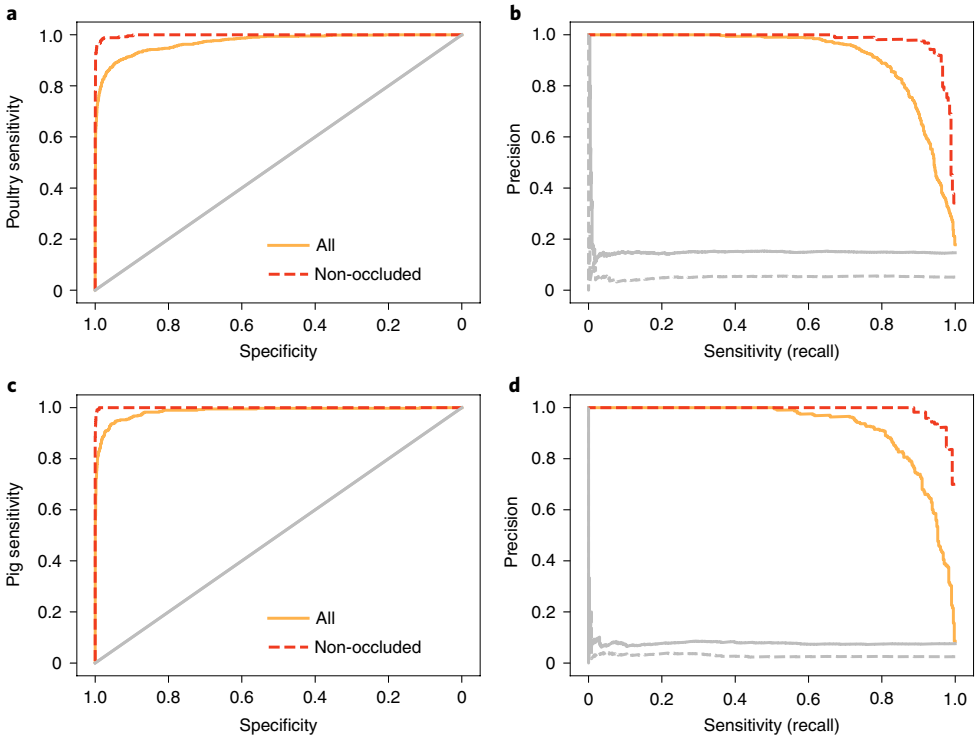

Fig. 2 | Classification performance. a–d, ROC curves (**a**,**c**) and precision-recall curves (**b**,**d**) for the poultry (**a**,**b**) and pig (**c**,**d**) models on the test set of images, split by whether the test sample included occluded CAFO images. The grey lines in each panel represent the performance of a random classifier predicting the presence of a poultry (**a**,**b**) or a pig (**c**,**d**) CAFO (dashed grey lines are associated with non-occluded images, solid grey lines with all images).

ing. Image tiles present a convenient unit of analysis but are limited in their ability to depict entire facilities due to the fact that facilities can span multiple images. Because tiles at times occluded substantial parts of the CAFO, we report accuracy metrics using (1) all CAFO and control images and (2) non-occluded CAFO images and all control images. We define 'non-occluded' as pig CAFO images with at least 25% of the image taken up by the CAFO facility or lagoon and, because poultry CAFOs tend to lack lagoons, poultry images with at least 10% of the image taken up by the CAFO facility (Fig. 1). Second, we report results from our facility identification, which addresses occlusion (when the facility is only partially in the image) by recentring images after an initial round of scoring and assigning latitude and longitude locations for poultry facilities within the images.

Image level. Because the optimal classification threshold (the probability level at which to classify an image as containing a CAFO) will vary on the basis of the end-user's tolerance for false-positive and false-negative images, we assessed performance at all classification thresholds between 0 and 1 for a holistic view of model performance. We first generated three standard classification accuracy metrics:

$$\text{Sensitivity} = \frac{\text{No. correct CAFO predictions}}{\text{No. CAFO images}}$$

$$\text{Specificity} = \frac{\text{No. correct control predictions}}{\text{No. control images}}$$

$$\text{Precision} = \frac{\text{No. correct CAFO predictions}}{\text{No. all CAFO predictions}}$$

where a CAFO prediction is correct if a CAFO image receives a CAFO probability score above a certain threshold, a control prediction is correct if a control image receives a CAFO score below that same threshold. Sensitivity (also known as recall) captures the proportion of correct predictions among CAFO images; specificity captures the proportion of correct predictions among control images; and precision, a measure of search relevance, captures the proportion of CAFO predictions that are truly CAFOs. The summary statistics that we present incorporate these metrics at all possible classification thresholds between 0 and 1.

With these metrics calculated at every classification threshold, we computed the area under the curve for the receiver-operating characteristic (ROC) curve, which plots specificity against sensitivity at each threshold and the precision-recall curve, which plots precision against sensitivity at each threshold. The area under the ROC curve provides a sense of average sensitivity across all possible values of specificity, whereas the area under the precision-recall curve provides a sense of average precision at all possible values of sensitivity. We report both sets of statistics because the precision-recall curve is recommended with class imbalance[30]. Relative to all of North Carolina, CAFOs are rare. On the basis of the EIG data, CAFO images constitute about 0.37% of the 1,684,879 images in North Carolina. Imbalance is more exaggerated statewide than in the test set owing to our attempt to manually validate as many CAFO images as possible. To provide a sense of performance on all images statewide, Supplementary Fig. 8 presents a weighted version of these metrics, using the weighting procedure described in the Supplementary Notes. For all metrics, values close to 1 indicate higher accuracy.

Table 1 presents the area under the curve values for the pig and poultry models, and Fig. 2 plots the corresponding ROC and precision-recall curves for all images (solid) and images without occlusion (dashed). Among all test images, the area under the curve for the poultry ROC curve is 0.97, which rises to over 0.99 for poultry and pig when excluding occluded CAFO images. We see comparable improvement in precision-recall curves when excluding occluded CAFO images. Table 1, however, shows that non-occluded

Fig. 3 | Illustration of class activation map algorithm for image-level classification. a, The algorithm first performs class activation mapping on the original image. **b**, Subsequently, the algorithm identifies the centroid of the activated clusters found by *k*-means clustering. **c**, The image is then recentred. **d**, Finally, the algorithm performs class activation mapping on the recentred image with both the pig and poultry models to identify the final class and size of the CAFO object. The poultry and pig scores in the top right corners of **b** and **d** indicate the predicted probability of the image containing a poultry or pig CAFO, respectively, obtained from the models.

CAFO images comprise only about 30% of all CAFO images, suggesting that more work is needed to transform image predictions into a useful format.

Facility level. As a substantive matter, environmental groups and regulators are more concerned with the locations of individual facilities than with the images. As is evident from Table 1, image occlusion stemming from the tile system may hinder detection of facilities. Figure 3 shows that when lagoons are occluded, a pig CAFO may appear close to a poultry CAFO, because the tile obscured the liquid manure storage system that is prototypical of pig CAFOs.

We therefore developed an approach to consolidate image-level predictions into latitude and longitude coordinates of unique facilities. We utilized 'class activation maps,' which depict which pixels activate the predicted class[31]. After classifying poultry CAFOs at the image level by taking the maximum score across the pig and poultry models, we used these activation maps to recentre the image on the CAFO. We then rescored the recentred images with both CNN models, and discarded any facilities for which the maximum score was no longer the poultry score or did not meet the poultry classification threshold. Finally, we clustered the activated pixels in the recentred image and represented the facility as a polygon shape, using the polygon centre as the point location. The end result is a list of latitudes and longitudes for (predicted) poultry facilities. Figure 3 depicts this process for a CAFO initially classified as poultry due to occlusion, but that was (correctly) reclassified as pig after recentring revealed the manure lagoon (see Supplementary Methods for details).

To provide a sense of performance, we manually validated the point locations for 4,659 predicted poultry facilities using a score threshold of 0.5 for the initial image-level predictions. We chose this threshold for demonstration purposes because it maintained a high harmonic mean of sensitivity and precision in both the weighted and unweighted test image samples (Supplementary Fig. 7). In practice, however, researchers may use different classification thresholds to meet their tolerance for false-positive and false-negative images. We also matched each confirmed poultry facility location to the nearest EIG location to determine the overlap between our results and the manual census. We considered model-identified CAFO locations to be the same as EIG locations when they were within 250 m of each other. Supplementary Figure 11 shows that this distance threshold correctly matched EIG locations to modelled locations in 99% of cases in a validation sample.

Through our earlier manual-validation efforts, we discovered that roughly 1.4% of EIG facility locations did not locate an existing poultry CAFO in reference to our NAIP imagery. Adjusting for this source of noise, we found that 73% of our predicted poultry facility locations were truly poultry CAFOs (analogous to precision),

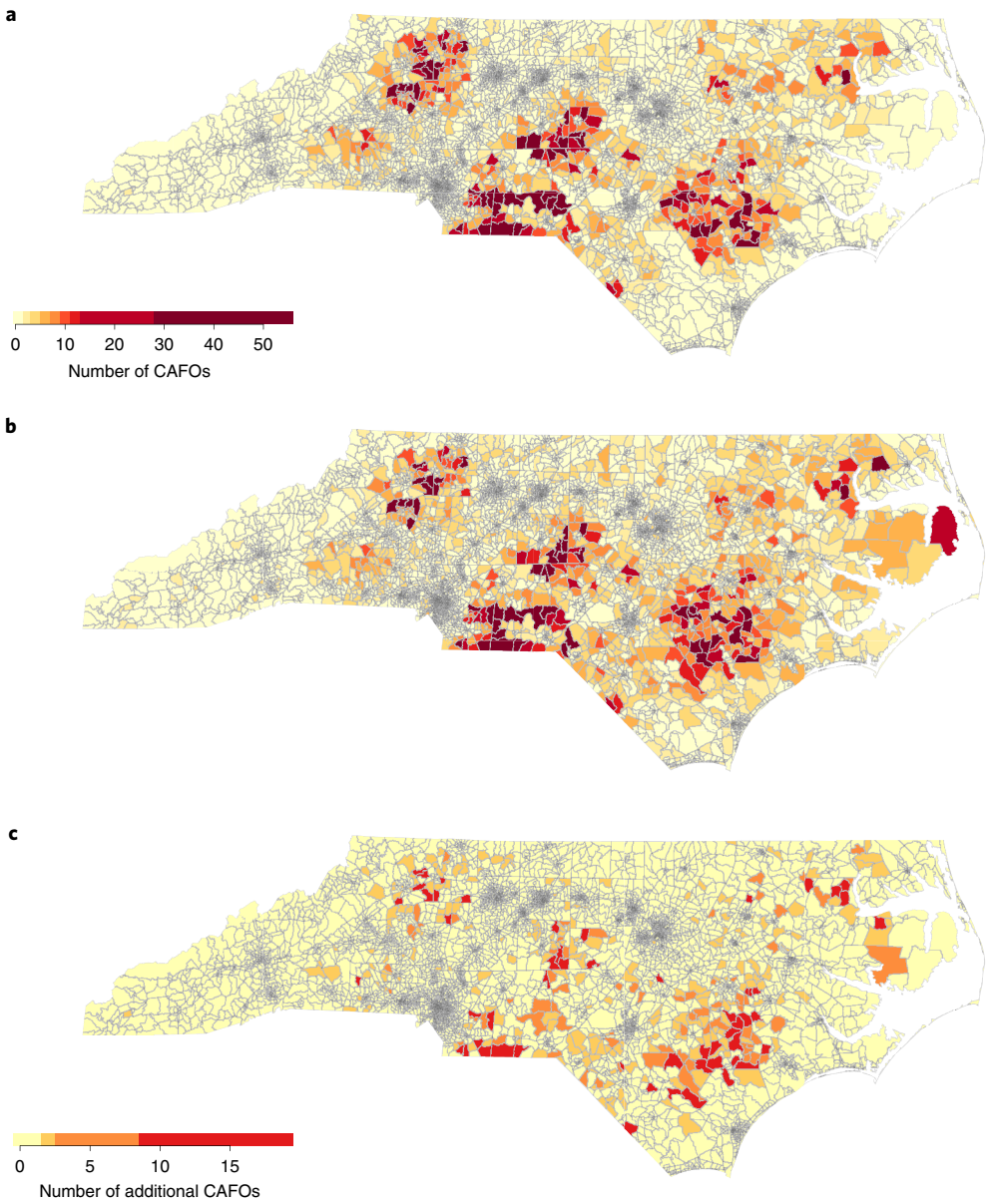

Fig. 4 | Heat map of manually identified and modelled poultry locations. a,**b**, The locations across North Carolina obtained by manual tagging efforts (**a**) and the fully automated machine-learning solution (**b**), including false-positive images. **c**, The 589 confirmed additional poultry facility locations.

whereas our predicted facility locations included 70% of EIG identified locations (analogous to sensitivity). It is worth noting that such precision is quite high given that CAFOs comprise only 0.4% of the raw image data. On average, modelled point locations were 60 m from the manually coded location provided by EIGs, when the length of a conventional CAFO barn can easily exceed 100 m.

Our model results and validation supplement EIG data in important ways. The validation efforts helped to identify difficult-to-classify facilities. Among the 1.4% of EIG locations that did not locate a poultry CAFO, for example, was a facility that visually appeared to be a CAFO, but was in fact a warehouse (Supplementary Fig. 13). Most importantly, although the model missed CAFOs that were manually detected by EIGs at a threshold of 0.5, it also detected an additional 589 CAFOs that EIGs were unable to detect. This is perhaps because the timing of the NAIP imagery (2014–2016) did not line up with the timing of the EIG manual census (2013–2014) or because, even to the human eye, poultry facilities can appear similar to other objects, such as airplane hangars and greenhouses, in aerial imagery because they lack outdoor lagoons. The newly detected facility locations represent a 15% gain from the manual census. Similar to capture–recapture methods[32], the model can help to identify facilities that may have been missed with the manual census of CAFOs.

One limitation to this facility-level analysis is that it is computationally intensive. That said, it is considerably less computationally demanding than image classification methods that use overlapping strides to ensure that CAFOs are centred in at least one image. Consolidating images into facility locations additionally reduces the burden on manual reviewers: at a threshold of 0.5, consolidation into facility locations reduces the number of items for review by nearly half (from 8,165 images to 4,659 facility locations).

Discussion

Having demonstrated the accuracy of a deep-learning approach, we illustrate some potential benefits and applications and conclude with future directions.

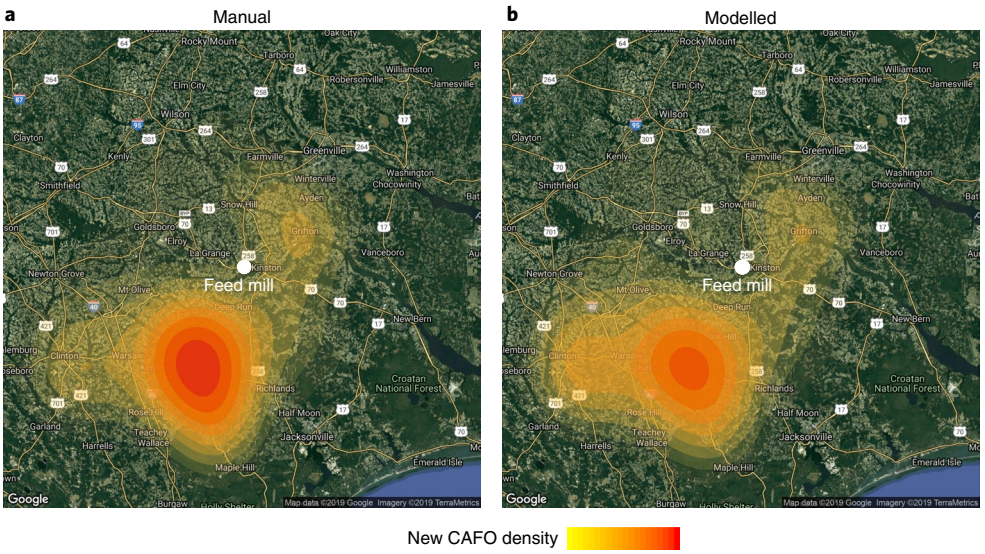

Fig. 5 | Longitudinal detection of CAFO growth. The density of poultry CAFO locations constructed after a feed mill opened in 2011 was quantified. **a**, Manually tagged density of CAFOs constructed after the feed mill opened. **b**, Modelled density.

Compared to the existing manual census, our approach promises to save considerable resources. For instance, the Environmental Working Group employs individuals to manually scan across a territory of satellite images on Google Maps, averaging 1.5 h per 1,000 km^2. To classify the entire state of North Carolina by manual tagging would require nearly 6 weeks of full-time work; to classify the entire United States would require over 7 years for a single imagery set. This manual census becomes infeasible as satellite coverage expands, for example, when including daily captures of the full globe[33]. By contrast, the automated approach can score North Carolina in fewer than 2 days with a standard workstation. The model may also be used as a complement to pinpoint where resources to validate images should be deployed. On the basis of the precision-recall curves weighted at the expected state-level distribution of images (269 control images for every CAFO image), we estimate that one could capture 75% of poultry CAFOs using less than 2% of manual resources or 95% of poultry CAFOs using less than 10% of manual resources (Supplementary Fig. 9). The facility consolidation approach reduces the use of manual resources even further—one can capture 70% of manually tagged poultry CAFOs using 0.28% of manual resources, while identifying an additional 15% of previously unknown facilities. Figure 4 provides a heat map of the distribution of poultry locations, comparing the manually identified locations with model-based locations. These maps show that our model is able to recover the geographical distribution of CAFOs extremely well.

Our approach also scales over time. This is important because the Government Accountability Office has expressly noted that the lack of consistent and reliable data has impeded the ability to understand changes in agricultural practices over time[13]. NAIP imagery data, however, were available at the same resolution in North Carolina from 2008 to 2016 (Supplementary Notes). Using these longitudinal data, we examined whether the construction of large industrial feed mills was associated with growth in poultry CAFOs[34]. We used public materials to identify the construction of feed mills in North Carolina and focused on one constructed in 2011. We applied our model to a 80-km radius around the feed mill, because this radius has been described to be the threshold for transportation costs[35], and manually tag historical NAIP images for validation. Figure 5 plots the area around the feed mill along with the density of poultry CAFOs that opened after the feed mill was constructed. The model detected 93% of all poultry CAFOs in the area and was 97% accurate in determining which ones appeared after the feed mill had opened.

Our approach may also help to determine NPDES compliance priorities. The class activation map can provide a meaningful measure for size, which is a critical dimension under federal law. The EPA has lamented the lack of information on size, as decentralized enforcement has resulted in missing, outdated or biased data, impeding prioritization across jurisdictions. We estimate the square footage of poultry CAFOs based on the activation maps (see Supplementary Methods for details). To assess the quality of estimates, we manually coded the number of barns for all model-detected facilities in Duplin and Cumberland counties. The Spearman rank correlation coefficient between the manually coded number of barns and the model size estimates is 0.46 ($P < 0.001$), showing that these estimates may be a meaningful way to enable environmental groups and regulators to prioritize resources.

In addition, our approach can facilitate the identification of facilities that pose particular risk due to proximity to water sources or abandonment. Supplementary Figures 14 and 15 provide specific examples. Abandoned CAFO sites may be cross-referenced with expired permit data to determine which facilities have undergone the appropriate clean-up procedures. Facilities operating under a Clean Water Act permit are required to remove waste and facilities from the site upon closure. Without such clean-up, the abandonment of all sites, permitted or not, can have considerable ecological consequences and should be prioritized for environmental compliance.

We conclude with some thoughts on future directions. First, although computer vision has been most rapidly adopted in the private sector, the public sector has been late to adopt artificial intelligence, creating a substantial 'technology gap'[36]. Our case study suggests that smart compliance, rather than displacing conventional enforcement[37], may free up the resources of nonprofit organizations and government agencies to focus efforts on facilities that pose the greatest risk of environmental harm. Partnerships between groups with the technical capacity and regulators are crucial to bridging the technology gap. Second, the challenge that we faced when trying to reconcile the EIG and NAIP data illustrates the importance of replicability. When nonprofit organizations or governments conduct exhaustive manual surveys, storing the underlying data (that is, image files or date stamps) should greatly facilitate the application of computer vision. Third, more work remains to be done to enable the efficient search

for, and detection and localization of, facilities. Last, although we demonstrated that our approach yields accurate predictions, prediction errors are of course still expected. Given that regulation demands high accuracy, a more appropriate solution may be a hybrid one, in which humans can iteratively refine model predictions[38].

We believe much work can be done to facilitate environmental monitoring, as satellite imagery data are available at continuously increasing resolution and frequency. One company, for instance, captures the globe's landmass on a daily basis[39] and another company captures images at 15 cm per pixel[40]. These improved data and the methods described herein may ultimately enable not only the identification of CAFOs, but also more genuine discharge measurements of CAFOs. Such applications may be useful not only in the context of the United States, but also in many other countries that have seen rapid growth in industrial agriculture[29].

Methods

Data. We started with two raw data sources. First, we obtained the most recent release of NAIP satellite images for North Carolina (2014–2016). We took advantage of the research platform from Descartes Labs, a geospatial analytics company, to download 299 × 299 image tiles, at 1 m per pixel, using the universal transverse mercator grid for North Carolina (Supplementary Notes). For the entire state, these raw NAIP data comprised 1,684,879 images (~32 gigabytes). Second, we obtained data from two environmental groups that contained point locations (latitude and longitude) of pig and poultry CAFOs from the manual census of North Carolina conducted between 2013 and 2014. The Environmental Working Group collected information on pigs and Waterkeeper Alliance collected information on poultry, which we call the EIG data. EIG point locations are based on permit data for pigs and a combination of ground monitoring and manual tagging of Google Earth satellite images for poultry. Owing to time differences in imagery data and imprecision in latitude and/or longitude (for example, referring to mailing address not agricultural facility), EIG point locations did not always resolve to a CAFO image in the NAIP data.

To resolve these differences in EIG and NAIP data for ground-truth data, we trained a team of student researchers to validate and tag NAIP images using a web application (Supplementary Fig. 1). For each image, coders were instructed to compare NAIP images against Google Maps and historical imagery on Google Earth where necessary, and to code animal type (pig, cattle, poultry or other) and facility features (building, manure storage type or feedlot). We validated our protocol for distinguishing CAFO types against the protocol independently derived by EIGs. Coders also indicated the proportion of the image depicting the CAFO and the proportion of the CAFO depicted in the image (see Supplementary Fig. 2 for examples). During the initial training of research assistants, we double-coded all images to ensure inter-coder reliability. Subsequently, the team met weekly to review and resolve classification conflicts on a randomly sampled subset of images.

We assigned the team (1) all NAIP images with an EIG (suspected CAFO) point location, (2) all images neighbouring an EIG point location on the universal transverse mercator grid (to match topology), and (3) randomly sampled control images, which were suspected not to contain a CAFO from the EIG census. Owing to resource constraints and the sheer number of control images, we developed a prototype model (98% accuracy) to help us to focus our efforts on more difficult control images to classify. Using scores from this model, we excluded control images with very low probabilities (below 0.007) of containing a CAFO from manual validation. We arrived at this score threshold through manual inspection of a random sample of results. Supplementary Figure 3 shows a representative set of excluded images, which tended to consist of trees and empty landscapes. The threshold excluded a total of 339,020 images out of 1,684,879 (about 20%) from eligibility to be sampled for manual validation. Supplementary Figure 4 confirms that these excluded images also had scores concentrated below 0.01 in the final poultry and pig models. We also oversampled difficult-to-classify images for manual validation by identifying Google's place categories that were associated with false-positive images in prototype models (see Supplementary Methods for details).

In total, the research team validated 24,440 images, with 3,385 and 1,599 images substantially depicting poultry and pig CAFOs, respectively. Interrater reliability for classification decisions was high (96% agreement).

Image modelling. We leveraged recent advances in image processing that have rapidly improved classification performance by CNNs, often dubbed 'deep learning'[41,42]. We applied transfer learning with Google's Inception V3 in the TensorFlow framework[43], which uses features learned on much larger image datasets, to retrain the final CNN layer for our classification problem[44].

Our classification problem involved three classes: poultry CAFOs, pig CAFOs and no CAFOs (control). There are several approaches to this type of multi-class problem. One approach would be to develop a single CNN to issue scores for all three classes at once. Another approach would be to develop one CNN to classify poultry CAFOs and another to classify pig CAFOs, with classification based on the higher of the two scores. Supplementary Figure 5 shows that the separate model approach outperformed the single model approach when it came to distinguishing poultry from pig images. Accordingly, we trained two separate CNNs to obtain probability scores for pig and poultry CAFOs independently.

We applied techniques common to CNNs to reduce overfitting on our training dataset[45,46]. In particular, we switched the left–right orientation of the images at random intervals, and randomly permuted the saturation (colour intensity) and brightness (contrast) of the RGB channels in the images. We also applied dropout to the input layer of our neural network, for which input units are randomly dropped during training with fixed probability to pare down the network[47]. Our classification problem involves substantial class imbalance: manual EIG coding produced 3,969 poultry locations and 2,291 pig locations out of 1,684,879 images of the entire state. Such class imbalance has been shown to negatively impact classification performance[48,49]. Accordingly, we naively oversampled images of CAFOs to match the number of images without CAFOs in the training set to enhance learning on the minority (CAFO) class. Supplementary Table 1 shows that naive oversampling outperformed other class-balancing techniques on prototype models. In addition, we oversampled difficult control images for manual validation. We randomly sampled high-scoring false-positive images from early rounds of model training and categorized them according to Google's place categories (see Supplementary Methods for details). These images were added to subsequent training rounds as control images. To optimize the learning rate (that is, the rate at which parameters of the loss function are updated in gradient descent), we used a triangular learning rate policy with an exponential decay parameter[50]. Supplementary Figure 6 provides a visualization of this procedure.

We randomly split our manually validated data into 60% for training, 15% for validation and 25% for testing. Because an image may only partially depict a CAFO (see Fig. 1 for examples of image occlusion), we experimented with limiting the CAFO images in the training set based on the fraction of the image depicting a CAFO and the fraction of the CAFO depicted in the image. Because pig CAFOs generally contain outdoor lagoons (manure storage pits), model performance improved when more of the image depicted the CAFO, increasing the chances that both facility and lagoon would be visible. Because poultry CAFOs tend to lack outdoor lagoons, more of the facility tended to fit within the image and performance improved when more of the CAFO was depicted in the image. Supplementary Table 2 describes the criteria that we used to limit CAFO images in the training set.

Facility detection. Because the image probability scores do not represent the proportion of a CAFO displayed in the image, image predictions do not provide the required facility locations (see the top panel of Supplementary Fig. 10 for an example). We required a localization approach to consolidate image-level predictions to a unique list of facilities with latitude and longitude point locations. The global average pooling layer of a CNN with Inception V3's architecture can perform localization tasks through a technique called class activation mapping[31]. This technique computes the dot product of the final softmax weights for a particular class and the feature map at the final global average pooling layer of the network to obtain a spatial representation of features that have activated the class.

By identifying the areas of the image that the model used to obtain the CAFO classification, class activation maps provide two additional pieces of information: where the activated pixels occur within the image, and a rough estimate of the object's size within the image. To provide the additional benefit of adding context to the object, we attempted to centre the CAFO within the image before using the class activation to estimate its size. By inspecting initial class activation maps, we saw that there may be many choices for the optimal centre of the facility. In addition, there may be several objects of interest within a single image. To be sure that we were exploring all the best possible centres identified by the model, we configured an algorithm to localize and compute the area of all possible CAFO objects within each image. The Supplementary Methods provide a detailed description of the procedure, and the bottom panel of Supplementary Fig. 10 provides a visual representation.

After running the algorithm described above on our true-positive poultry images, we scored all centred images once more with the pig and poultry models. This gave us the benefit of using newfound context after recentring to discard images that appeared to depict poultry facilities only because the liquid manure storage lagoon was occluded in the first instance (Fig. 3). Once the objects were confirmed as poultry by the second round of scoring, we developed a method to combine objects that represented the same CAFO within different images. First, we calculated the distances between all CAFO object centroids in metres on the universal transverse mercator grid. Next, we iteratively combined each object into a pair with its closest neighbouring object(s) if the neighbouring object's centroid was within 250 m of its centroid. The 250-m threshold was obtained by testing a range of thresholds and manually validating the accuracy of the resulting facility locations. We define the centroid for the newly paired set of objects as the average of all objects' centroids. As a size estimate, we keep the maximum square footage among all objects as the facility size, after prioritizing higher scoring objects and objects that are fully contained within the image. We repeated this process until the minimum distance between all CAFO objects was at least 250 m. In this way, we

reduced 9,601 possible CAFO objects statewide to 4,659 unique facilities. A detailed description on how we validated this approach can be found in the Supplementary Methods.

Data availability

The replication code and datasets generated during the current study are available in the GitHub repository github.com/slnader/cafo_public.

Received: 12 September 2018; Accepted: 15 February 2019;
Published online: 8 April 2019

References

1. Dimitri, C., Effland, A. & Conklin, N. *The 20th Century Transformation of U.S. Agriculture and Farm Policy* Economic Information Bulletin 3 (USDA, 2005).
2. Copeland, C. *Animal Waste and Water Quality: EPA Regulation of Concentrated Animal Feeding Operations (CAFOs)* Technical Report No. CRS-RL31851 (Congressional Research Service, 2010).
3. *40 C.F.R. 122.23 - Concentrated Animal Feeding Operations (Applicable to State NPDES Programs)* (Environmental Protection Agency, 2011); https://www.govinfo.gov/app/details/CFR-2011-title40-vol22/CFR-2011-title40-vol22-sec122-23
4. Casey, J. A., Kim, B. F., Larsen, J., Price, L. B. & Nachman, K. E. Industrial food animal production and community health. *Curr. Environ. Health Rep.* **2**, 259–271 (2015).
5. Hribar, C. *Understanding Concentrated Animal Feeding Operations and Their Impact on Communities* (National Association of Local Boards of Health, 2010).
6. Environmental Protection Agency National pollutant discharge elimination system permit regulation and effluent limitation guidelines and standards for concentrated animal feeding operations (CAFOs). *Fed. Regist.* **68**, 7176–7274 (2003).
7. Rogers, S. & Haines, J. *Detecting and Mitigating the Environmental Impact of Fecal Pathogens Originating from Confined Animal Feeding Operations* (EPA, 2005).
8. Graham, J. P. & Nachman, K. E. Managing waste from confined animal feeding operations in the United States: the need for sanitary reform. *J. Water Health* **8**, 646–670 (2010).
9. Conerly, O. & Vazquez Coriano, L. *Literature Review of Contaminants in Livestock and Poultry Manure and Implications for Water Quality* Report No. EPA 820-R-13-002 (EPA, 2013).
10. *Waterkeeper v. EPA* [2005] 399 F.3d 486 (2d Cir. 2005).
11. *Nat'l Pork Producers Council v. EPA* [2011] 635 F.3d 738 (5th Cir. 2011).
12. Environmental Protection Agency National pollutant discharge elimination system (NPDES) concentrated animal feeding operation (CAFO) reporting rule. *Fed. Regist.* **76**, 65431–65458 (2011).
13. *Concentrated Animal Feeding Operations: EPA Needs More Information and A Clearly Defined Strategy to Protect Air and Water Quality from Pollutants of Concern* Technical Report No. GAO-08-944 (US Government Accountability Office, 2008).
14. Environmental Protection Agency National pollutant discharge elimination system (NPDES) concentrated animal feeding operation (CAFO) reporting rule. *Fed. Regist.* **77**, 42679–42682 (2012).
15. *American Farm Bureau Fed. v. EPA* [2016] 836 F.3d 963 (8th Cir. 2016).
16. Brown, C. R. Uncooperative federalism, misguided textualism: the federal courts' mistaken hostility toward pre-discharge regulation of confined animal feeding operations under the clean water act. *Temple J. Sci. Technol. Environ. Law* **30**, 175–219 (2011).
17. Jerger, S. EPA's new CAFO land application requirements: an exercise in unsupervised self-monitoring. *Stanf. Environ. Law J.* **23**, 91–130 (2004).
18. Moses, A. & Tomaselli, P. in *International Farm Animal, Wildlife and Food Safety Law* (eds Steier, G. & Patel, K.) 185–214 (Springer, 2017).
19. Pew Commission on Industrial Farm Animal Production *Putting Meat on the Table: Industrial Farm Animal Production in America* (Pew Charitable Trusts and Johns Hopkins Bloomberg School of Public Health, 2007).
20. Peterka, A. Enviro groups return CAFO data at heart of Hill probes. *E&E Daily* https://www.eenews.net/stories/1059979265/print (11 April 2013).
21. Formuzis, A. Fields of filth: landmark report maps feces-laden hog and chicken operations in North Carolina. *Environmental Working Group* https://www.ewg.org/release/fields-filth-landmark-report-maps-feces-laden-hog-and-chicken-operations-north-carolina (22 June 2016).
22. *Factory Farm Nation 2015 Edition* (Food & Water Watch, 2015); https://www.factoryfarmmap.org/wp-content/uploads/2015/05/FoodandWaterWatchFactoryFarmFinalReportNationMay2015.pdf
23. *Stopping CAFO Pollution* (Michigan Chapter, Sierra Club, 2018); https://www.sierraclub.org/michigan/stopping-cafo-pollution
24. *Complaint Under Title VI of the Civil Rights Act of 1964* (EarthJustice, 2014); https://go.nature.com/2VRyw15
25. *2017 Annual Report for Work Plan Agreement Between the Iowa Department of Natural Resources and the Environmental Protection Agency Region 7* (Iowa Department of Natural Resources, 2017).
26. Kamilaris, A. & Prenafeta-Boldú, F. X. Deep learning in agriculture: a survey. *Comput. Electron. Agric.* **147**, 70–90 (2018).
27. Osburn, C. L., Handsel, L. T., Peirels, B. L. & Paerl, H. W. Predicting source of dissolved organic nitrogen to an estuary from an agro-urban coastal watershed. *Environ. Sci. Technol.* **50**, 8473–8484 (2016).
28. Patt, H. *A Comparison of PAN and P_2O_5 Produced from Poultry, Swine and Cattle Operations in North Carolina* (Division of Water Resources, North Carolina Department of Environmental Quality, 2017); https://go.nature.com/2NWY3Df
29. Ilea, R. C. Intensive livestock farming: global trends, increased environmental concerns, and ethical solutions. *J. Agric. Environ. Ethics* **22**, 153–167 (2009).
30. Saito, T. & Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE* **10**, e0118432 (2015).
31. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. & Torralba, A. Learning deep features for discriminative localization. In *IEEE Conference on Computer Vision and Pattern Recognition* 2921–2929 (IEEE, 2016).
32. Amstrup, S. C., McDonald, T. L. & Manly, B. F. *Handbook of Capture–Recapture Analysis* (Princeton Univ. Press, 2010).
33. Hurst, N. How daily images of the entire Earth will change the way we look at it. *Smithsonian* https://www.smithsonianmag.com/innovation/how-daily-images-entire-earth-will-change-way-we-look-it-180962467/ (13 March 2017).
34. MacDonald, J. M. et al. *Contracts, Markets, and Prices: Organizing the Production and Use of Agricultural Commodities* Agricultural Economic Report No. aer-837 (US Department of Agriculture, 2004).
35. McVey, M. J. et al. Identifying an efficient feed distribution system in the Midwest. In *SAEA 2003 Annual Meeting* https://go.nature.com/2Cece28 (Southern Agricultural Economics Association, 2003).
36. Newcombe, T. Is government ready for AI? *Government Technology* http://www.govtech.com/products/Is-Government-Ready-for-AI.html (1 July 2018).
37. Walton, B. EPA turns away from CAFO water pollution. *Circle of Blue* https://www.circleofblue.org/2016/water-policy-politics/epa-turns-away-cafo-water-pollution/ (2016).
38. Branson, S. et al. Visual recognition with humans in the loop. In *European Conference on Computer Vision* Vol. 6314 (eds Daniilidis, K. et al.) 438–451 (Springer, 2010).
39. Scoles, S. 88 new satellites will watch Earth, all the time, all the places. *Wired* https://www.wired.com/2017/02/88-tiny-satellites-will-watch-time-everywhere/ (14 February 2017).
40. *What are the technical specifications for Google Imagery?* (Google, 2018); https://support.google.com/mapsdata/answer/6261838?hl=en
41. Russakovsky, O. et al. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vision* **115**, 211–252 (2014).
42. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *3rd International Conference on Learning Representations* https://go.nature.com/2TwYLwN (Computational and Biological Learning Society, 2014).
43. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. In *IEEE Conference on Computer Vision and Pattern Recognition* 2818–2826 (IEEE, 2016).
44. Oquab, M., Bottou, L., Laptev, I. & Sivic, J. Learning and transferring mid-level image representations using convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition* 1717–1724 (IEEE, 2014).
45. Simard, P. Y., Steinkraus, D. & Platt, J. C. Best practices for convolutional neural networks applied to visual document analysis. 7th International Conference on Document Analysis and Recognition. Institute of Electrical and Electronics Engineers. 958-962 (2003).
46. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* Vol. 25 (eds Pereira, F. et al.) 1097–1105 (ACM, 2012).
47. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
48. Buda, M., Maki, A. & Mazurowski, M. A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.* **106**, 249–259 (2018).
49. Mazurowski, M. A. et al. Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance. *Neural Netw.* **21**, 427–436 (2008).
50. Smith, L. N. Cyclical learning rates for training neural networks. 2017 IEEE Winter Conference on Applications of Computer Vision. Institute of Electrical and Electronics Engineers. 464-472 (2017).

Acknowledgements

We thank Z. Ashwood, G. Hong, C. Hull and A. Teuscher for research assistance, the Environmental Working Group for sharing data, Descartes Labs and Google Earth Engine for providing access to their research platform, the GRACE Communications Foundation and the Stanford Institute for Economic and Policy Research for generous support, and C. Cox, B. Erden, S. A. C. Gomez, M. Hancher, M. Engelman Lado, J. Lee, P. Lehner, D. Lobell, J. Quinlivan, S. Rundquist, D. Sivas and the participants of the Data for Sustainable Development class at Stanford University for helpful conversations.

Author contributions

C.H.-N. and D.E.H. jointly designed the study, collected data, developed the methods, performed the analysis and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41893-019-0246-x.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to D.E.H.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.