Deep Learning with Satellite Imagery to Enhance Environmental Enforcement

Data-Driven Insights and Decisions: A Sustainability Perspective
(Jennifer B. Dunn & Prasanna Balaprakash eds. 2020, forthcoming)

Cassandra Handan-Nader
Stanford University

Daniel E. Ho
Stanford University

Larry Y. Liu
Stanford University

Abstract

In this chapter we highlight how rapid advances in computer vision and the increasing availability of high-resolution satellite imagery have facilitated more accurate, efficient, and scalable environmental monitoring and regulation. First, we highlight the range of potential use cases of remote sensing with satellite imagery in environmental enforcement. Second, we describe the methodological evolution from manual learning from satellite imagery, to model-based inference largely based on pixel-by-pixel classification, to deep learning. Third, we provide an in-depth case study, illustrating how deep learning with satellite imagery can solve a problem that has vexed the Environmental Protection Agency for decades: the identification of Concentrated Animal Feeding Operations (CAFO), which pose substantial environmental risk. Last, we highlight the data infrastructure, modeling, and capacity challenges that must be overcome to facilitate this profound shift in the evidence base for environmental enforcement.

I. Introduction **3**

II. The Methodological Evolution of Remote Sensing **7**

II.A. Manual Image Classification 8

II.B. Automatic Image Classification 10

II.C. Deep Learning 13

III. Using Deep Learning to Identify CAFOs **16**

III.A. Data 16

III.B. Transfer Learning 18

III.C. Results 21

IV. IV. Discussion **26**

IV.A. Challenges and Opportunities 27

IV.B. Future Use Cases 29

I. Introduction

The protection of air, water, and land depends critically on the role of government agencies that monitor and enforce environmental laws. In the United States, the Environmental Protection Agency administers a vast range of statutory schemes,[1] with regulations touching on critical industries, including energy, agriculture, transportation, and construction. Notwithstanding landmark statutes, such as the Clean Air Act and the Clean Water Act, there is increasing evidence that regulatory bodies struggle in enforcing these laws (Evans & Malcom, 2019; GAO, 2008, 2009; Salomon, Markus, & Dross, 2014).

We argue that the vast increase in the quantity and quality of satellite imagery, coupled with rapid advances in computer vision, often dubbed the “deep learning” revolution, has the potential to substantially enhance environmental monitoring and enforcement. Satellite imagery is often characterized by three features: (1) spatial resolution (e.g., meters per pixel), (2) temporal resolution (e.g., how often and when the satellite scans the same region), and (3) spectral sensitivity (e.g., a sensor’s dynamic range and measurement along the electromagnetic spectrum beyond basic colors (red, green, and blue)). The use of spectral characteristics has been particularly useful in the processing of satellite imagery, as the reflectance of objects -- the ratio of intensity of light reflected from a surface divided by the intensity of the incident light -- can provide information about land cover, vegetation characteristics, topography, surface temperature and precipitation, atmospheric properties, and water (Horning, 2008).

Consider the potential use cases for satellite imagery:

---

[1] For instance, the Environmental Protection Agency bears substantial responsibility for administering the Clean Air Act, the Clean Water Act (or Federal Water Pollution Control Amendments of 1972), the Comprehensive Environmental Response, Compensation and Liability Act (or Superfund), the Emergency Planning and Community Right-to-Know Act, the Federal Insecticide, Fungicide, and Rodenticide Act, the Safe Drinking Water Act, and the Toxic Substances Control Act.

- Montana regulators consulted Google Earth to detect zoning violations by buildings (Puckett, 2014).
- Massachusetts environmental authorities manually compared satellite imagery against permit records and identified 3,000 locations of unpermitted wetland filling by (Clayton, 2004).
- Researchers developed a method to use satellite imagery to detect habitat changes (e.g., oil and gas development) threatening biodiversity (Evans & Malcom, 2019).
- A Silicon Valley firm promises to use satellite imagery to track carbon emissions from specific power plants (Morris, 2019).

These examples may only scratch the surface for environmental protection (Onoda & Young, 2017). We review here a selected number of prominent examples for water, land, and air.

1. Water. A leading example of the use of satellite imagery for water monitoring is the detection of oil spills, an application that has gained rapid traction with regulators. Some early examples include a satellite-based oil spill alarm system in Norway (Wahl et al., 1996) and a pilot project to develop low cost detection of oil spills for environmental regulators in Spain (Martínez & Moreno, 1996). In 2002, the Canadian Space Agency (CSA) started funding the Integrated Satellite Tracking of Pollution project to detect oil spills, which has since been validated and integrated into enforcement operations. The process identifies suspected oil spills from satellites and feeds this information to surveillance aircraft in real-time. Such evidence has successfully been used in court to prosecute suspected illegal oil dumpers (Gauthier, Weir, Ou, Arkett, & Abreu, 2007). CleanSeaNet has been monitoring for potential oil spills on behalf of 27 European nations and five beneficiary countries since 2007 (Ray Purdy, Harris, Carver, &

Slater, 2017). Brekke & Solberg (2005) describe several methods for oil spill detection that can search large areas even at night and through cloud cover. Between June 2003 and March 2004, researchers identified 274 spills across 230 images (400x400 km, 75 meters / pixel resolution) in the South Baltic Sea (Kostianoy et al., 2006).

2. Land. Many satellite-assisted land monitoring efforts have focused on habitat fragmentation and deforestation. Early informative projects used satellite imagery to construct historical estimates of deforestation and habitat fragmentation in the Brazilian Amazon (Skole & Tucker, 1993) and Madagascar's eastern rainforests (Green & Sussman, 1990). The National Institute for Space Research now has several satellite-based early warning detection systems for deforestation in the Brazilian Amazon, with experimental systems drawing on satellite images at time intervals of just 5 days (Diniz et al., 2015). Australia has used aerial photography to demonstrate illegal land clearance over time, offering such evidence in enforcement suits (R. Purdy, 2010). Many European countries use satellite data to assist regulators in determining compliance with the pre-conditions for agricultural subsidy programs (Ray Purdy et al., 2017).

3. Air. Satellites have increased the precision of air pollution monitoring by measuring the concentration of certain compounds in the atmosphere. NASA's Geostationary Carbon Cycle Observatory measures atmospheric carbon dioxide (CO_2), carbon monoxide (CO), and methane (CH_4) with a horizontal ground resolution of 3-6 miles. Satellites can also measure nitrogen dioxide (NO_2), sulfur dioxide (SO_2), ammonia (NH_3), volatile organic compounds, aerosol optical depth, allowing for estimation of surface particulate matter ($PM_{2.5}$) (Duncan et al., 2014). Witte et al. (2009) used satellite data to study China's emissions curtailments to improve air quality for the 2008 Olympics. They found large and significant reductions in air pollution, including 43% decrease in tropospheric column nitrogen dioxide in Beijing, demonstrating the efficacy of these regulations in improving air quality. Commercial satellites promise to measure greenhouse gases from individual facilities (Aganaba-Jeanty & Huggins, 2019).

Satellites can also fill gaps in air pollution sensors. For instance, only 21 percent of counties have sensors to monitor particulate matter ($PM_{2.5}$) performance under the National Ambient Air Quality Standards. Sullivan & Krupnick (2018) used satellite data to estimate that roughly 24 million people reside in 54 counties across 11 states that are misclassified with ground monitors for particulate matter. Such sensors exhibited false negative rates exceeding 50%. Li, Shen, Yuan, Zhang, and Zhang (2017) combined satellite, ground, and meteorological data to develop better estimates of particulate matter.

* * * *

These existing use cases illustrate the promise of satellite data in environmental monitoring and enforcement. Indeed, one research team concluded that a significant number of the 106 EU laws and 42 international laws could be monitored by satellites (R. Purdy, 2010). Notwithstanding this potential and these prominent examples, “the actual application of [satellite imagery in] environmental compliance [remains] more theoretical than applied” (id.). There are two major reasons to believe that such use cases are likely to grow significantly. The first is that is that the number of active satellites has increased dramatically over the past 10 years. Figure 1 plots year on the *x*-axis against the number of active satellites on the *y*-axis, showing a sharp rise in the last decade, with some 2,062 registered satellites currently orbiting the Earth (Union of Concerned Scientists, 2019). The sheer growth in satellite fleets has also been accompanied by dramatic improvements in the quality of imagery. In the 1970s through the 1990s, the spatial resolution of satellites was around 30-80 meters per pixel. Current commercial satellites produce images at less than .50 meter (R. Purdy, 2010). The second reason for increased reliance on satellite imagery in environmental monitoring and enforcement is the one we turn to in Section 2: rapid methodological advances in computer vision. These new methods have enhanced our capacity to synthesize the increasingly large amounts of data from satellites into output that humans can readily make sense of and put to use for environmental regulation.

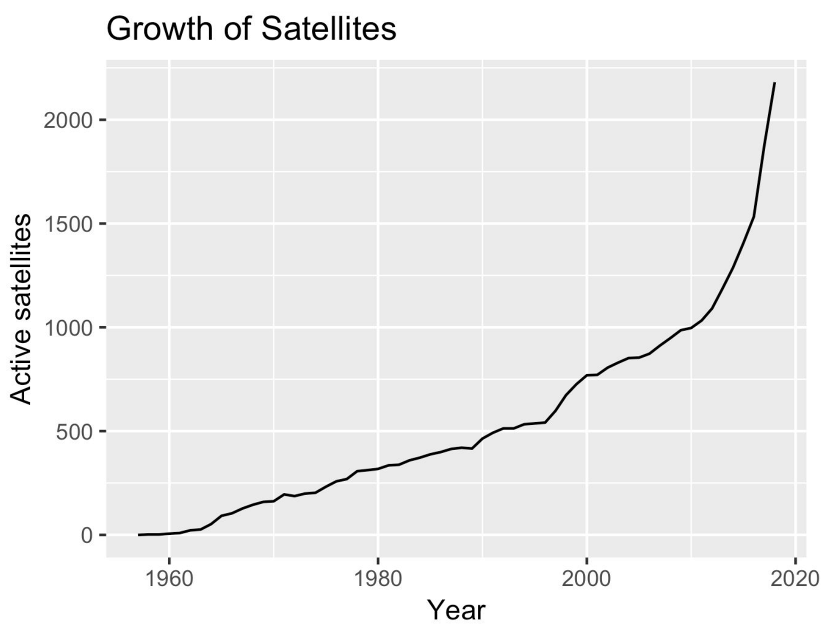

Figure 1: Growth in satellite fleet over time. This figure plots the number of active satellites on the *y*-axis against year on the *x*-axis. Source: Statistica.

II. The Methodological Evolution of Remote Sensing

We now review the evolution of image processing methods in remote sensing, with a particular focus on satellite imagery. These methods have rapidly evolved from manual learning, to automated and model-based image processing, and most recently, to deep learning. Our review highlights connections between often disparate literatures in remote sensing, computer vision, and machine learning, and we argue that many advances will come from cross-fertilization between these fields.

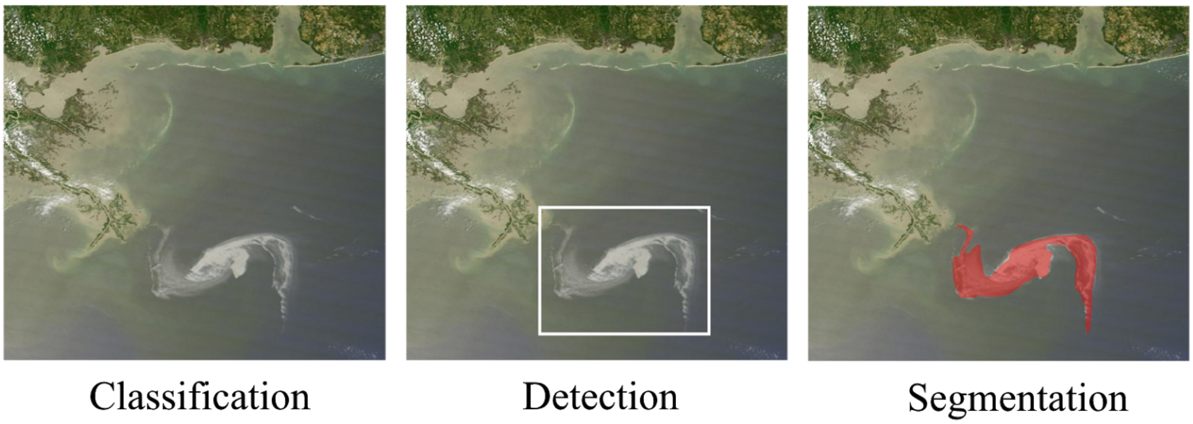

Figure 2: Types of image learning tasks with Gulf Oil Spill image. The left panel depicts the task of classifying an image tile as containing an oil spill. The middle panel depicts the task of localizing the oil spill by drawing a bounding box around it. The right panel depicts the task of segmenting individual pixels into whether they depict an oil slick or not. Source: NASA, [MODIS image]() of Gulf Oil Spill from July 29, 2010.

Many tasks in remote sensing and computer vision boil down to identifying whether an object of interest is present in an image. One can accomplish this task at varying levels of granularity. Image classification, the least granular approach, involves determining whether an image contains an object of interest anywhere in its boundaries (e.g., whether an image contains an oil spill). A more granular approach known in computer vision as object detection aims to draw a box around the object of interest within the image (e.g., a "bounding box" around an oil spill within an image). The most granular is image segmentation, which involves drawing a polygon around the object of interest, hence identifying the particular pixels comprising the object (e.g., whether pixels depict an oil slick). Figure 2 displays these different image tasks with oil slicks.

Much of remote sensing has approached these tasks by extracting lower-level features to identify patterns relating to the larger object of interest. This approach tracks the history of computer vision, that also started by building images up level by level.[2] Conventionally, remote sensing methodology has focused on pixel-by-pixel classification, such as classifying a pixel into land cover type. As a result, a major challenge to this approach has accounting for the spatial context of a set of pixels within the image as a whole. Computer vision has recently taken a more distinct path. With the "deep learning" revolution, computer vision has turned towards learning lower-level features based on their ability to aid with the ultimate task (e.g., image classification). In contrast, much of remote sensing continues to conceive of feature creation (i.e., pixel classification) as distinct step before focusing on objects that comprise multiple pixels. For instance, conditional on land cover classification of each pixel, a remote sensing method might attempt to construct higher-level features from such given labels to detect an object within an image ("object-based image analysis" (T. Blaschke, 2010)). Features in the object-based image analysis are determined ex ante based on pixel attributes, and are not iteratively updated based on how well they perform in the ultimate image learning task, as they are in deep learning models.

Below we trace the evolution of image processing in remote sensing. First, we spell out remote sensing via manual image processing in Section II.A. We then turn to automated image processing, including maximum likelihood classification and machine learning techniques largely deployed for pixel classification by convention in Section III.B. In Section III.C, we describe the turn to deep learning.

[2] This conception of computer vision as dividing into different levels of analysis is often attributed to David Marr (Glennerster, 2007).

II.A. Manual Image Processing

Manual methods of reviewing image data obtained from remote sensing technologies have been thoroughly investigated over the past half century (Hollings et al., 2018). When classifying images manually, an analyst extracts or interprets key information from the imagery data by evaluating color, tone, texture, size, shape, shadow and context (Slonecker, Jennings, & Garofalo, 2001).

Löffler & Margules (1980), for instance, used Landsat satellite images to monitor the spread of wombats thought to be responsible for destroying significant areas of shrubland in Australia. Relying on field knowledge of wombat habitation behavior and the appearance of their colonies, the authors noticed conspicuous white circular areas on black and white Landsat images. Subsequent aerial images and field visits confirmed that these areas corresponded to concentrated wombat warrens. Thus, based on iterative field- and manual visual inspection, the authors concluded that "black and white imagery provides a tool to map the approximate distribution of wombat colonies."

Manual methods, however, are often time-consuming and require the analyst to be familiar with the area covered by the satellite image. The efficiency and accuracy of the classification depend on the analyst's expertise in the field of study, and human subjectivity can generate inconsistencies in pattern and interpretation (Brodrick, Davies, & Asner, 2019). That said, manual review remains an important element of existing environmental enforcement systems. Automated oil spill detection systems, for instance, still trigger surveillance aircraft to verify that an actual oil spill has occurred.

II.B. Automated Image Processing

Automatic approaches offer an alternative to time-consuming, manual inspection, and remote sensing researchers have applied a range of approaches. These methods include unsupervised learning, rule-based "expert systems," parametric classifiers, and machine learning. A principal tension in this literature has been how to move from a primary focus on pixel-by-pixel classification to detection and classification of objects that span more than a single pixel (T. Blaschke, 2010). Efforts in remote sensing to move to "object-based" or contextual approaches are similar in aim to object detection in computer vision, which explains the advantages of convolutional neural networks, discussed in Section II.C below.

1. Unsupervised Learning. The earliest applications -- still in use today -- corresponded to the pattern recognition movement in computer vision and forms of unsupervised learning. Rouse et al. (1973) first described the Normalized Vegetation Difference Index (NVDI), used to approximate the presence and density of live green vegetation. NVDI is defined as:

$$NVDI = (NIR - Red)/(NIR + Red),$$

where NIR and red are the spectral reflectance measurements acquired in the near-infrared and red regions, respectively.[3] NVDI values around -1 highly suggest the presence of water; around 0, the presence of bare soil; and around +1, dense green leaves. Other normalized spectral difference indices have since been developed (Gao, 1996).

[3] Healthy vegetation reflects a large portion of near-infrared light, while bare soils by contrast reflect moderately in both the red and near-infrared portions of the electromagnetic spectrum. The difference between near-infrared and red reflectance values corresponds to the amount of vegetation. The larger the difference, the greater the vegetation. The denominator of the NVDI normalizes this difference, accounting for instances where different (NIR – Red) values could be obtained for two identical patches of vegetation under different cloud cover and sunlight conditions.

Compared to a hand-engineered formula, clustering is a more data-driven approach to group segregate spectrally similar pixels, with an analyst often attaching a meaningful class label after clustering. Clustering has been applied to change detection for deforestation or seasonal changes in agricultural production (Ashbindu, 1989). For a more recent application, Leichtle, Geiß, Wurm, Lakes, & Taubenböck (2017) apply clustering to monitor dynamic urban environments.

2. Expert-Based Systems. In the 1970s, artificial intelligence researchers began to investigate knowledge-intensive "expert systems"—logical AI systems that represent expert judgments as a set of Boolean decision rules (Russell & Norvig, 2016). Expert systems primarily seek to reproduce human expertise in algorithmic form (Goodenough, Goldberg, Plunkett, & Zelek, 1987), and they proliferated in the 1980s, including in computer vision and remote sensing.[4] For instance, Goodenough et al. (1987) described the "Landsat Digital Image Analysis System (LDIAS)," intended to support the analysis of an image into up to 32 classes in eight hours. Taking some eight years of development and boasting more than one million lines of Fortran code, LDIAS required a significant amount of hardware to implement—three VAX computers,

[4] Matsuyama (1989), for instance, describes a "Low Level Vision Expert" (LLVE) system, which incorporated know-how about image segmentation into seven "production rules," used to accomplish tasks such as extracting line segments from a gray picture. A typical analysis process for this task in LLVE would be: Gray Picture (Edge Detection) → Edge Picture (Thresholding) → Edge Point (Linking) → Line Segment. Gray Picture, Edge Picture, Edge Point, and Line Segment are *image features*, or information extractable from raw image data. Edge Detection, Thresholding, and Linking are *transfer processes*, or abstract algorithms that analyze an input image feature and generate an output image feature. Each transfer process is associated with a group of executable image processing operators. The Edge Detection transfer process, for example, is associated with various edge detection filters such as the Sobel and Laplacian filters. LLVE decides which transfer process to apply, in what sequence, the specific operator or filter, and the operator or filter's corresponding parameters. The quality of these decisions is ultimately tied to the amount and quality of knowledge represented by the production rules.

three AI VAXstations, four image displays, two map displays, and several special processors.[5] Matsuyama & Hwang (1990) developed the Sigma system for aerial imagery, using a "blackboard model" that represents domain knowledge about each object (e.g., a house, road). Over the course of years, expert-based systems continue to be developed. Forestier, Puissant, Wemmert, & Gançarski (2012) built a domain-specific knowledge base to allow for region-based labeling of segmented images. Yet as LDIAS illustrates, the expert-based approach has proven challenging to operationalize at scale and across domains (Russell & Norvig, 2016).

3. Supervised Learning. Supervised learning, where researchers train models based on datasets labeled with ground truth, have proven more fruitful to deploy at scale. In remote sensing, the predominant approach has been supervised learning with pixel-by-pixel classification. The canonical remote sensing application has been land cover classification (e.g., labeling a pixel as grassland, forest, or urban). We review some of the dominant methods in remote sensing.

<u>Maximum Likelihood Classification (Multivariate Normal Class)</u>. While statisticians will speak of maximum likelihood estimation as a general approach to fit a wide range of models (including neural networks), the remote sensing literature tends to use the term to refer to the simpler multivariate normal class model (Richards, 2013). We follow that terminology below. Hixson, Scholz, Fuhs, & Akiyama (1979) described several maximum likelihood classifiers to identify major crops (e.g., corn, soybeans, and small grains) in the mid-Western United States. Training data were used to estimate parameters for the multivariate normal distribution governing the generation of each class. Multivariate normal class models remain frequently used, and for many applications, remain the dominant method (Richards, 2013). In a review article evaluating

[5] VAX (Virtual Address Extension) was an instruction set architecture designed in the mid 1970s to better execute programs written in high-level programming languages. VAXstation was a family of workstation computers using processors implementing the VAX instruction set architecture.

land cover classification, Yu et al. (2014) noted that such maximum likelihood classifiers were the modal method, used in 32.34% of 1,651 studies.

This type of maximum likelihood classification remains popular in spite of substantial research adapting machine learning methods for remote sensing that demonstrate higher accuracy (Maxwell, Warner, & Fang, 2018; Richards, 2013). Yu et al. conjecture that its popularity stems from widespread availability in conventional remote-sensing image-processing software packages. Maxwell et al. (2018) similarly conjecture that the principal barrier for using machine-learning methods is uncertainty about their use and implementation by applied researchers. We hence turn to some of the most widely used machine learning methods in remote sensing.

Support Vector Machines (SVMs). In contrast to the multivariate normal classifier described above, most machine learning methods disregard the probability distribution of the data and simply focus on maximizing separation between a predefined number of classes. One example of this approach known as SVMs has become increasingly popular in remote sensing (Belgiu & Drăguț, 2016). SVMs use training data to maximize the margin between data points defining the hyperplane that separates the classes (Mountrakis, Im, & Ogole, 2011). Mountrakis et al. identify 108 remote sensing articles deploying SVMs in 2.5 years prior to their review, with about an equal split between application and methodology. SVMs have been used in various land cover and land use remote sensing tasks such as vegetation and crop classification, as well as evaluating urban areas and detecting impervious surfaces (Mountrakis et al., 2011).

Decision Trees and Random Forests. Another machine learning method focused on class separation is the decision tree. Decision trees are flowchart-like structures that represent how attributes (e.g., spectral characteristics of an image pixel) at each “node” lead to different class label probabilities (e.g., water, vegetation, land) (Richards, 2013). In contrast to expert-based systems, which encode expert knowledge, the nodes of a decision tree are learned from training data. Random forests constitute a particularly powerful ensemble technique that grows a large number of trees, with the classification decision averaging the class assignment probabilities across trees (Breiman, 2001). Random forests have been applied to mapping land cover classes, classifying impervious surfaces, and identifying oil spills (Belgiu & Drăguţ, 2016).

Artificial Neural Networks (ANNs). ANNs are often conceptualized as a mathematical analogue to the brain's axons and their connections through synapses (Atkinson & Tatnall, 1997; Maxwell et al., 2018). ANNs "learn" by dynamically adjusting the weights associated with each neuron in the network from training data, a process known as back propagation. The first ANNs were applied in remote sensing in the early 1990s. Hepner et al. (1990) fit (by current standards a simple) ANN (one hidden layer with 10 perceptrons) to classify pixels of an image of the Ft. Lewis Military Reservation near Tacoma, Washington into four land cover classes (water, built, forest, and grass). Foreshadowing developments in convolutional neural networks, the ANN used a 3-by-3 pixel and four-band (RGB and near infrared) array as an input to enable incorporation of information of spatially adjacent pixels, reducing the so-called "salt-and-pepper" effect with pixel-by-pixel classification (T. Blaschke, Lang, Lorup, Strobl, & Zeil, 2000). They found that the ANN architecture both required less training data and outperformed a multivariate normal class model (maximum likelihood).

4. Object-Based Image Analysis. One of the research frontiers emerging in the 2000s was the transition from pixel to object-based analysis (T. Blaschke, 2010; Thomas Blaschke et al., 2014). Conventional maximum likelihood classifies each individual pixel, and object-based image analysis has aimed to move beyond pixels to objects comprised of many pixels (e.g., roads, buildings, parks). The approach starts with pixel classes, segments images based on pixel classes (e.g., grouping by polygon shape), and then conducts object classification based on the spectral characteristics of the object. One of the first commercial programs to offer such object-based analysis (eCognition) has most recently turned to deep learning techniques, signaling the change in image classification.

II.C. Deep Learning

Most recently, *deep learning* models have beat conventional image learning benchmarks. In contrast to the simple ANN used by Hepner (1990), deep learning neural networks use convolutions of image features in many layers (hence, the name “convolutional neural network” or CNN). CNNs start from pixel inputs and progressively abstract more complex features, such as shapes and edges, in order to make a final prediction of the image class. More specifically, CNNs learn custom filters that, when convolved with the image data in one hidden layer, produce higher-level features as new inputs for subsequent hidden layers of the network. Periodically, the convolved features pass through pooling layers to reduce dimensionality. In this way, CNNs can capture spatial relationships among the image pixels that correlate with the image class. In a classification model, the final layer performs a “softmax” transformation to calculate a predicted probability for each class.

Two developments made the modern deep learning revolution possible. First, the development of affordable graphics processing units (GPUs) in the early 2000s enabled fitting complex CNNs. Initially designed for dedicated graphics rendering for computer games, GPUs have been adopted by data scientists because of their ability to rapidly compute matrix and vector operations. In neural network training, using GPUs can speed up learning by a factor of 50 or more compared to using central processing units (CPUs) (Schmidhuber, 2015). Modern

GPU-based computers are a million-times computationally more powerful than the desktops of the 1990s (Schmidhuber, 2015).

Second, the development of large labeled datasets enabled the training of CNNs. Early datasets of labeled images such as NORB (LeCun, Fu Jie Huang, & Bottou, 2004), Caltech-101/256 (Li Fei-Fei, Fergus, & Perona, 2004), and CIFAR-10/100 (Krizhevsky, 2009), were relatively small, comprising tens of thousands of images. In 2009, ImageNet's release heralded a paradigm shift within machine learning. ImageNet contained over 15 million labeled high-resolution images in over 22,000 categories. For the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC), Krizhevsky, Sutskever, & Hinton (2012) famously submitted a convolutional neural network (CNN) architecture that dominated its competition—outperforming the second-highest scoring submission in classification accuracy by 10.8%. This "AlexNet" model ushered in a new wave of interest in CNNs.

Figure 3 shows that in the last five years, the field of deep learning has exhibited exponential growth, from non-existence in 2012 to over three thousand papers published on the topic by this year.

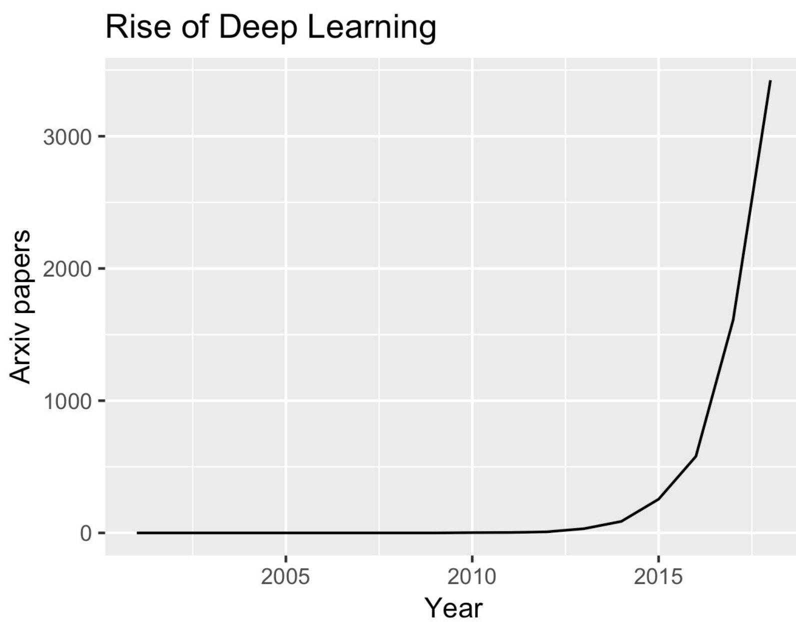

Figure 3: Rise of deep learning over time. This figure plots the number of Arxiv papers with the term “deep learning” on the *y*-axis against year on the *x*-axis. Results are comparable when plotting proportion of Arxiv papers with the term “deep learning.”

Although deep learning and remote sensing remain distinct literatures, some scholars have begun to apply to CNNs remotely sensed images (Brodrick et al., 2019). Längkvist, Kiselev, Alirezaie, & Loutfi (2016) demonstrate how CNNs can improve both segmentation and object classification of satellite imagery. Rezaee et al. (2018) applied a CNN to optical satellite images to classify wetlands land cover in a 700 km^2 area in Newfoundland and Labrador, Canada. And CNNs have also recently been applied to satellite images to classify land cover (Syrris et al., 2019), estimate animal populations (Guirado, Tabik, L. Rivas, Alcaraz-Segura, & Herrera, 2018; Yifei Xue, Tiejun Wang, & Andrew K. Skidmore, 2017), detect oil tanks (Wang, Zhang, Hu, & Wang, 2016), and predict poverty (Jean et al., 2016; Xie, Jean, Burke, Lobell, & Ermon, 2016).

The above examples, however, are not representative of the field. A recent review of machine learning for remote sensing noted that “[m]achine-learning classification has become a major focus of the remote-sensing literature.” But that review “[did] not consider . . . deep convolution[al] neural networks, because such methods have not yet been widely adopted” (Maxwell, Warner, & Fang, 2018). Part of the reason for the lack of uptake may be that training a full CNN can be costly in terms of data and computation. The emergence of “transfer learning,” however, eases those burdens substantially making deep learning much more widely available (Oquab, Bottou, Laptev, & Sivic, 2014). Transfer learning adapts existing image models by retraining only the last few layers to the domain-specific task. This allows researchers to borrow image features from pre-trained models (e.g., AlexNet) trained on large image datasets (e.g., ImageNet), hence lowering the requisite training sample size. Hu, Xia, Hu, & Zhang (2015) demonstrate the potential for transfer learning with remote sensing. They use standard image models (e.g., AlexNet) and either (a) retrain the last (fully connected) layer or (b) extract features from intermediate convolutional layers, and demonstrate that such transfer learning improves performance with with classifying scenes (e.g., airport vs. farmland).

* * * *

The modern state of the art reflects parallel advances in our ability to capture images at high resolution and to store and process that data. As Slonecker et al. (2001) put it, “the science of remote sensing is currently undergoing a dramatic revolution in terms of data type and availability” promising to provide “a new paradigm” in the imaging and analysis of environmental phenomena. Methods have evolved from visual review of image features to model-based pixel classification, to machine-learning, to object-based detection, and most recently to deep-learning for image classification. To illustrate a deep learning model in more detail, we turn now to a case study involving the application of a CNN to detect concentrated feed animal operations.

III. Using Deep Learning to Identify CAFOs

Concentrated animal feeding operations (CAFOs) are estimated to produce nearly 40% of U.S. livestock (Copeland, 2010) and generate some 13-25 times the manure of humans. But animal waste is not required to be treated, and poses considerable environmental and health risks to water, air, and land (Conerly & Vasquez Coriano, 2013; Graham & Nachman, 2010; Rodgers & Haines, 2005). Due in part to substantial litigation, no consistent, reliable public data source exists documenting the size and location of these operations (GAO, 2008). Geographically proximate CAFOs that house similar types of animals tend to share similar visual characteristics, and, with some training, a human could quickly learn how to pick out a set of CAFOs from aerial satellite imagery. As such, detecting CAFOs seems a well-suited task for satellite imagery in combination with the methodologies discussed in the previous section.

Yet the task has proven to be a difficult one for traditional remote sensing techniques. In 2004, the EPA pioneered an effort to identify the locations of swine CAFOs within a delimited area of Duplin County, North Carolina, by triangulating pixel-based land cover classifications with hand-collected measurements on distances between hog barns and manure lagoons (Garofalo & Jennings, 2004). However, land cover classification methods are severely limited by the fact that CAFOs, which do not cultivate crops yet can be surrounded by cropland, often defy traditional land-use categories. Martin, Emanuel, & Vose (2018) cross-referenced manually collected CAFO location data in North Carolina with land cover classes from the US National Land Cover Database (NLCD) and found that a full 57% of permitted and 35% of non-permitted CAFOs would be classified as cultivated cropland by the NLCD. Perhaps worse, 27% of permitted CAFO locations and 8% of non-permitted locations would be classified as natural systems by the NLCD (e.g., forest, wetlands). A subsequent effort to use rules-based object analysis on pixel-level values in tandem with automated feature detection proved more successful within the same delimited area of Duplin County (Feingold, Zaitchik, & Silbergeld, 2011), but the generalizability and scalability of the approach to other areas of the state remained unexplored.

In 2011, when the EPA considered implementing more comprehensive reporting requirements for CAFOs, it mentioned the use of satellite imagery as an unsolved but promising supplement to on-the-ground approaches (EPA, 2011). The proposed rule also mentioned several ongoing data collection processes to incorporate satellite imagery into monitoring efforts, including flyovers by the EPA itself. But despite the promise of early efforts and the expansion of data collection, relatively few open-source, scalable prototypes for detecting CAFOs using satellite imagery have emerged in the years following the report.[6] In this section, we show that a form of deep learning called transfer learning may provide a more flexible, efficient, and scalable approach to this task than traditional remote sensing techniques (Handan-Nader & Ho, 2019).[7]

III.A. Data

To build out our proof of concept, we required high resolution satellite imagery and training data to develop a model. Our approach was image classification, meaning that the training data consisted of labels associated with tiles of satellite imagery. We now describe the steps to generate this training data.

First, we rely on the U.S. Department of Agriculture's National Agricultural Imagery Program (NAIP), which provides comprehensive imagery across the United States at relatively

[6] One preliminary effort similar in spirit to the approach described in this section is documented at https://github.com/Qberto/ML_ObjectDetection_CAFO.
[7] Our discussion here draws significantly on Handan-Nader & Ho (2019).

high resolution over a relatively long range of time. Since 2005, NAIP has been available at resolutions of 1-2 meters per pixel across the continental US on a three-year cycle. Though the spatial resolution of NAIP is quite high, it lacks the temporal resolution necessary to measure changes at a daily or even yearly level on a consistent basis. In addition, it has historically only provided the red-green-blue (RGB) spectral bands, though it recently began providing a near-infrared band.

Second, we secured CAFO locations (latitude, longitude, animal type) from a manual enumeration conducted by two environmental interest groups in North Carolina from 2013-14. We hence focus on North Carolina, which is home to a large number of hog and poultry CAFOs and has been the subject of much controversy (Formuzis, 2016; Nicole, 2013).

Third, to make the raw satellite imagery tractable for modeling, we divided imagery into tiles using the Universal Transverse Mercator (UTM) grid system. Descartes Labs, a research platform, made it easy to download NAIP imagery in 299 x 299 image tiles for the entire state of North Carolina, with 6 pixels of overlap across images. (We chose this 299 x 299 tiling system, as it allowed us to use pretrained image models, explained below.) At 1 meter per pixel, these image dimensions captured a large proportion of a typical CAFO facility while remaining a computationally feasible input size for modeling. The image data consisted of 1,684,879 image tiles, or about 32 gigabytes of data.

Fourth, because the most proximate NAIP images were taken between 2014-16, we trained a research team to manually validate the environmental interest group data against NAIP imagery and developed a system to handle such tagging of images dynamically.[8] As our

[8] We sampled all images whose bounding boxes contained a CAFO found by the environmental groups, and a random sample of all other images in the state. Once we had built a few high-functioning prototype models, we excluded from this latter group of images those that had very low predicted probabilities of containing CAFOs. In this way, we focused the attention of our research team primarily on images suspected to contain CAFOs.

objective was to classify images into ones that either contained (a) poultry CAFOs, (b) swine CAFOs, or (c) no CAFOs (control images), we trained the research team to recognize features of each type of CAFO. We developed our training materials independently from the environmental groups, then validated our process against theirs. Our final labeled dataset contained a little over twenty-four thousand images. We reserved a quarter of this dataset for testing of the model predictions, and the remaining for training and validation.

III.B. Modeling Approach

Transfer Learning. One of the most practically useful developments in computer vision has been the rise of "transfer learning." Because CNNs take the raw image pixels as input and learn all the filters from scratch, they typically require training data on the order of hundreds of thousands (if not millions) of images to achieve high accuracy (Krizhevsky et al., 2012). As described earlier, collecting training data was manually intensive work for the environmental groups, and there is necessarily a limited number of CAFO images within a given geographic area. Transfer learning substantially reduces training data requirements by using a CNN trained on a very large image corpus. Transfer learning uses these parameters and higher-level image features to re-train a final softmax layer to predict new image classes. Recent research has shown that transfer learning can work well in other remote sensing applications (Xie et al., 2016) as well as more conventional image recognition tasks (Oquab, Bottou, Laptev, & Sivic, 2014).

Preprocessing. We took several pre-processing steps to maximize the effectiveness of model development. First, we performed standard "data augmentation" procedures to increase the variety of our training data, such as randomly flipping images and permuting their color intensity and saturation. This endeavor was particularly important for the models' generalizability to longitudinal applications, as NAIP images can vary in color profile over time. Because CAFO images were relatively rare across all image tiles (about 1 in 270), we also applied class balancing techniques to improve accuracy (Buda, Maki, & Mazurowski, 2018). We found that naively oversampling images from the CAFO classes to match the number of control images improved accuracy the most. Finally, we used Google Places to locate difficult-to-classify "trick images," such as mobile home parks and airplane hangars, and supplemented the control image pool with these images.

Classification Models. Several approaches could have been adopted to classify CAFOs with animal types. We experimented with multilabel classification, but because images could feasibly contain more than one type of CAFO, we found that developing two separate models to predict the swine and poultry classes against the control class yielded better results than training a single model that requires predicted probabilities for swine, poultry, and control classes to sum to 1. These models provided two separate scores between 0 and 1 for each image that represented the probability of the image containing a swine or poultry CAFO. Having two separate models also allowed us to independently analyze poultry CAFOs, which is of particular interest given the laxer permitting regime for poultry in North Carolina.[9]

Facility Identification. As a practical matter, identifying facilities is of greater interest than classifying image tiles. In addition, estimates of facility size would be useful to set enforcement priorities. We hence took additional steps to (a) *localize* the facility within images predicted to contain a poultry CAFO, (b) map that location back to geographic coordinates on the UTM grid and centering the facility within a new image tile, and (c) provide a rough square footage estimate for the facility. We did this using the "class activation maps" derived from our final CNN models (Zhou, Khosla, Lapedriza, Oliva, & Torralba, 2016). These maps consist of the dot product of the final softmax weights for a particular class and the final feature map at the

[9] This is because the use of a dry litter storage system does not require a separate permit.

last pooling layer. Intuitively, the map conveys which areas of the image have “activated” the predicted class. Figure 4 illustrates a class activation map and how our localization and recentering allows for improved classification of animal type.

Figure 4: Example of how class activation maps can help reveal the location and type of CAFO in an image. (a) The original image shows a CAFO facility that could be either swine or poultry. (b) The class activation maps highlight where the CAFO facility is located within the image. (c) We recenter the image on the area highlighted by the class activation map. (d) Once the liquid manure storage is contained within the image, the swine score exceeds the poultry score and the image is reclassified as swine.

III.C. Results

Classification accuracy on the test set of images was quite high, particularly when limiting the CAFO images to only those that showed a sizable portion of the facility (e.g., AUC = 0.99 for poultry[10]). Performance dropped slightly, however, when including images in which the

[10] AUC refers to the area under the receiver operating characteristic curve.

CAFO facility was substantially cut off, as the model lacked context outside of the image boundaries. Nonetheless, when weighting the images to approximate the true distribution of CAFO images across the entire state, we estimated that this model performance could save considerable resources in the attempt to capture all the known CAFOs in the state through manual enumeration. For instance, to capture 95% of all known CAFOs in the state, one could tag fewer than 10% of the images than were tagged by environmental interest groups by simply reviewing the highest scored images returned by the model.

To assess model performance at the facility level, we manually validated 4,659 predicted poultry facility locations. We found that 73% of these were true poultry facilities. Since CAFO images comprised only 0.4% of the raw image data, this figure represents a large improvement over an exhaustive manual search of images. Not only did our facility list include 70% of the locations that the environmental groups manually identified, it also included 15% more locations not included in the manual census, likely because of timing differences in the two efforts. Since the facility consolidation process further reduces the number of items for a human to review, these figures suggest that model-assisted detection can capture 70% of previously located poultry CAFOs using only 0.28% of manual resources, while detecting 15% more facilities than previously known. The model's facility size estimates were meaningful, correlating significantly with the number of barns in the image (Spearman rank correlation coefficient = 0.46, $p < 0.001$). Figure 5 illustrates the geographic distribution of poultry CAFOs obtained by the manual census, the modeled census, and the new facilities found by the model. We found that this modeling approach scales temporally as well as geographically. We applied our poultry model to a set of images from 2008 to 2016 that were within a 50-mile radius of a feed mill constructed in 2011. The model was able to detect which facilities appeared after the construction of the feed mill with 97% accuracy.

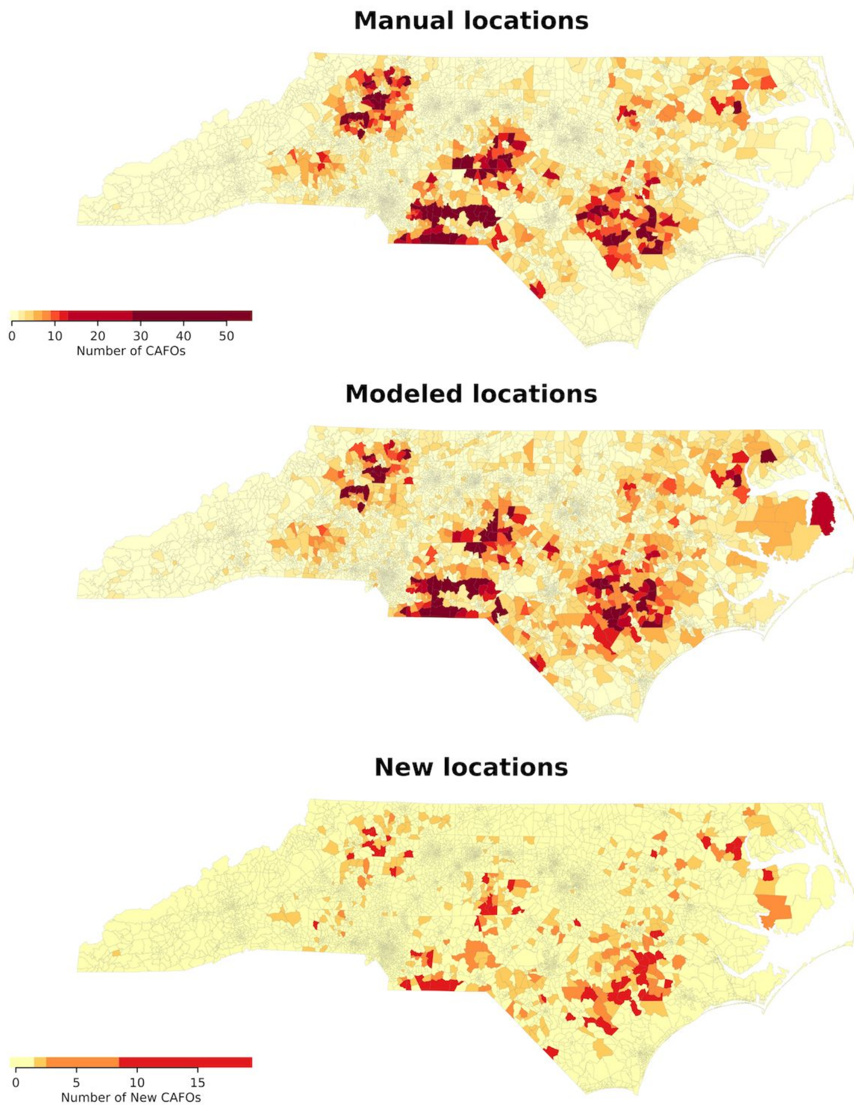

Figure 5: Heatmap depicting poultry CAFO counts from the manual census (top), the fully automated modeled census, including false positives (middle), and the new locations found by the model (bottom).

Our case study demonstrates how a lightweight transfer learning approach can approximate the geographic distribution of poultry CAFOs with a high degree of accuracy. Rather than serving as a replacement for human-driven efforts, these models can free up human resources to focus on more complex monitoring efforts by surfacing areas of interest to scrutinize more carefully. Of course, model predictions will always include some degree of error. In practice, humans can participate in mitigating common sources of model error by iteratively reviewing predictions (Branson et al., 2010). Model accuracy would also likely improve with higher resolution satellite imagery, which is becoming more widely available in commercial settings.

IV. Discussion

We close with thoughts on challenges that must be overcome to facilitate such usage of satellite data on a broader scale in environmental enforcement.

Data. Our in-depth use case illustrates several challenges with data infrastructure. First, while satellite imagery has rapidly grown, there remains a dearth of ground truth training data, particularly localized object labels, compared to other image domains (Van Etten, 2018). Recent advances to create more benchmark datasets will be critical for growth in this field (Liu, Yuan, Weng, & Yang, 2017; Mundhenk, Konjevod, Sakla, & Boakye, 2016; Van Etten, Lindenbaum, & Bacastow, 2018). Second, more thought needs to be given to unify data standards to enable linkages across datasets. Despite the availability of labels and satellite imagery in our setting, the time disparity required comprehensively re-validating image labels. Third, our work was only possible due to large-scale publicly available NAIP imagery. If new image data is only available on a commercial basis, much of the public good in environmental applications may be lost. In recent years, the USDA, for instance, has considered making NAIP imagery available only on a fee basis (Mootz, 2017), which would significantly constrain the prototyping of solutions for environmental enforcement.

Modeling. Rapid platform advances such as Descartes Labs and Google Earth Engine have made processing of satellite imagery substantially easier. That said, the modeling pipeline is not yet as advanced for satellite imagery as it is for conventional imagery. First, leading approaches have not yet addressed the unique computational complexities associated with processing large-scale satellite imagery. More work needs to adapt architectures to efficiently scan over vast volumes of satellite imagery, such as with those proposed by Van Etten (2018, 2019). Second, while transfer learning from conventional image models worked well for our application, few pre-trained models using satellite imagery exist, and such models could enable more rapid transfer learning with smaller training datasets (see, e.g., Jean et al., 2018; Lu, Zheng, & Yuan, 2017). Third, existing methods that have largely been developed for visible (RGB) spectrum need to be developed to incorporate the wider range of information across multiple satellites across the electromagnetic spectrum (Audebert, Le Saux, & Lefèvre, 2018). We believe that greater exchange between previously distinct remote sensing and machine learning communities may be particularly valuable for advancing methods for learning with satellite imagery.

Inference and Evaluation. While advanced machine learning has rapidly advanced our capacity to learn from satellite imagery, it is also worth noting the limitations of machine learning. Most importantly, these methods are not a substitute for statistical methods to draw an inference about causal effects. Deep learning with satellite imagery over time may inform our understanding of ecological changes, but cannot directly allow us to draw causal inferences about interventions. This is particularly important in the domain of environmental enforcement, the legal or policy question of interest may be (a) whether an actor caused environmental harm, or (b) whether the automated detection system and/or the marginal enforcement action based on satellite imagery (e.g., warning letter, site inspection, educational visit) warrant the cost. While computer vision with satellite imagery can improve the information basis for environmental monitoring, systems that use such information should be piloted, tested, and evaluated.

Agency Capacity. When it comes to the ultimate adoption of these techniques by government agencies, growth is constrained due to (a) legacy software and database systems, (b) human capital, and (c) the institutional setting for adopting rapid advances in machine learning (Cuellar, Engstrom, Ho, & Sharkey, 2019). More generally, some have noted a kind of cultural impediment, namely a "skepticism regarding the operational utility and reliability of remotely sensed data as an environmental compliance tool" (Lein, 2009).[11] To overcome that skepticism requires rigorous piloting, testing, and evaluation, as well as a recognition that effective enforcement requires triaging of resources, with early steps not necessarily meeting 100% accuracy. One promising avenue is that some environmental agencies have begun to collaborate with academic institutions. Coming out of its NextGen Compliance plan, for instance, EPA has worked with several academics to bring insights from cutting edge research into enforcement and compliance strategies.[12]

* * * *

While these challenges are substantial, overcoming them will be critical to leveraging cutting edge technology and data science. If these obstacles can be tackled as in other domains of computer vision, these developments promise to modernize environmental enforcement to secure compliance across a wide range of domains more effectively, accurately, and efficiently.

---

[11] As an extreme example, responding to national attention, one town passed an ordinance "restricting use of satellite imagery to conduct 'sweeps' in place of field inspections and investigations" (Knoedler, 2012).

[12] Stanford University's Regulation, Evaluation, and Governance Lab and the University of Chicago's Energy and Environment Lab have partnered with environmental agencies to pilot data science in environmental enforcement.

References

Aganaba-Jeanty, T., & Huggins, A. (2019). Satellite Measurement of GHG Emissions: Prospects for Enhancing Transparency and Answerability under International Law. *Transnational Environmental Law*, *8*(2), 303–326. https://doi.org/10.1017/S2047102519000104

Ashbindu, S. (1989). Digital change detection techniques using remotely-sensed data. *Int. J. Remote Sensing*, *10*(6), 989–1003.

Atkinson, P. M., & Tatnall, A. R. L. (1997). Introduction Neural networks in remote sensing. *International Journal of Remote Sensing*, *18*(4), 699–709. https://doi.org/10.1080/014311697218700

Audebert, N., Le Saux, B., & Lefèvre, S. (2018). Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. *Geospatial Computer Vision*, *140*, 20–32. https://doi.org/10.1016/j.isprsjprs.2017.11.011

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004

Blaschke, T., Lang, S., Lorup, E., Strobl, J., & Zeil, P. (2000). Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In *Environmental Information for Planning, Politics and the Public* (Vol. 2, pp. 555–570). Marburg: Metropolis Verlag.

Blaschke, Thomas, Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., … Tiede, D. (2014). Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of*

Photogrammetry and Remote Sensing, *87*, 180–191. https://doi.org/10.1016/j.isprsjprs.2013.09.014

Branson, S., Wah, C., Schroff, F., Babenko, B., Welinder, P., Perona, P., & Belongie, S. (2010). Visual Recognition with Humans in the Loop. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010* (pp. 438–451). Springer Berlin Heidelberg.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brekke, C., & Solberg, A. H. S. (2005). Oil spill detection by satellite remote sensing. *Remote Sensing of Environment*, *95*(1), 1–13. https://doi.org/10.1016/j.rse.2004.11.015

Brodrick, P. G., Davies, A. B., & Asner, G. P. (2019). Uncovering Ecological Patterns with Convolutional Neural Networks. *Trends in Ecology & Evolution*, S0169534719300862. https://doi.org/10.1016/j.tree.2019.03.006

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, *106*, 249–259. https://doi.org/10.1016/j.neunet.2018.07.011

Clayton, M. (2004, January 22). Wetlands get a new spy cam. *Christian Science Monitor*. Retrieved from https://www.csmonitor.com/2004/0122/p11s02-sten.html

Conerly, O., & Vasquez Coriano, L. (2013). *Literature Review of Contaminants in Livestock and Poultry Manure and Implications for Water Quality* (No. EPA 820-R-13-002). Environmental Protection Agency.

Copeland, C. (2010). *Animal Waste and Water Quality: EPA Regulation of Concentrated Animal Feeding Operations (CAFOs)* (No. CRS-RL31851). Congressional Research Service.

Cuellar, M.-F., Engstrom, D. F., Ho, D. E., & Sharkey, C. (2019). *Administering by Algorithm: Artificial Intelligence in the Regulatory State* [Administrative Conference of the United

States].

Diniz, C. G., Souza, A. A. d A., Santos, D. C., Dias, M. C., Luz, N. C. d, Moraes, D. R. V. d, … Adami, M. (2015). DETER-B: The New Amazon Near Real-Time Deforestation Detection System. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(7), 3619–3628. https://doi.org/10.1109/JSTARS.2015.2437075

Duncan, B. N., Prados, A. I., Lamsal, L. N., Liu, Y., Streets, D. G., Gupta, P., … Ziemba, L. D. (2014). Satellite data of atmospheric pollution for U.S. air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. *Atmospheric Environment*, *94*, 647–662. https://doi.org/10.1016/j.atmosenv.2014.05.061

EPA. (2011). National Pollutant Discharge Elimination System (NPDES) Concentrated Animal Feeding Operation (CAFO) Reporting Rule. *Federal Register*, *76*(204), 65431–65458.

Evans, M. J., & Malcom, J. W. (2019). Automated Habitat Change Detection Methods using Satellite Data to Improve Conservation Law Implementation. *BioRxiv*, 611459. https://doi.org/10.1101/611459

Forestier, G., Puissant, A., Wemmert, C., & Gançarski, P. (2012). Knowledge-based region labeling for remote sensing image interpretation. *Computers, Environment and Urban Systems*, *36*(5), 470–480. https://doi.org/10.1016/j.compenvurbsys.2012.01.003

Formuzis, A. (2016). *Fields of Filth: Landmark Report Maps Feces-Laden Hog and Chicken Operations in North Carolina*. Retrieved from https://www.ewg.org/release/fields-filth-landmark-report-maps-feces-laden-hog-and-chicken-operations-north-carolina

GAO. (2008). *Concentrated animal feeding operations: EPA needs more information and a clearly defined strategy to protect air and water quality from pollutants of concern* (No.

GAO-08-944). US Government Accountability Office.

GAO. (2009). *The US Fish and Wildlife Service Has Incomplete Information about Effects on Listed Species from Section 7 Consultations*. GAO-09-550.

Gauthier, M., Weir, L., Ou, Z., Arkett, M., & Abreu, R. D. (2007). Integrated satellite tracking of pollution: A new operational program. *2007 IEEE International Geoscience and Remote Sensing Symposium*, 967–970. https://doi.org/10.1109/IGARSS.2007.4422960

Glennerster, A. (2007). Marr's vision: Twenty-five years on. *Current Biology*, *17*(11), R397–R399. https://doi.org/10.1016/j.cub.2007.03.035

Goodenough, D. G., Goldberg, M., Plunkett, G., & Zelek, J. (1987). An Expert System for Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-25*(3), 349–359. https://doi.org/10.1109/TGRS.1987.289805

Graham, J. P., & Nachman, K., E. (2010). Managing waste from confined animal feeding operations in the United States: The need for sanitary reform. *Journal of Water and Health*, *8*(4), 646–670.

Green, G. M., & Sussman, R. W. (1990). Deforestation History of the Eastern Rain Forests of Madagascar from Satellite Images. *Science*, *248*(4952), 212–215. https://doi.org/10.1126/science.248.4952.212

Guirado, E., Tabik, S., L. Rivas, M., Alcaraz-Segura, D., & Herrera, F. (2018). *Automatic whale counting in satellite images with deep learning* [Preprint]. https://doi.org/10.1101/443671

Handan-Nader, C., & Ho, D. E. (2019). Deep learning to map concentrated animal feeding operations. *Nature Sustainability*, *2*(4), 298–306. https://doi.org/10.1038/s41893-019-0246-x

Hixson, M. M., Scholz, D., Fuhs, N. N., & Akiyama, T. (1979). *Evaluation of several schemes for classification of remotely sensed data: Their parameters and performance. [Foster*

County, North Dakota; Grant County, Kansas; Iroquois County, Illinois, Tippecanoe County, Indiana; and Pottawattamie and Shelby Counties, Iowa].

Hollings, T., Burgman, M., van Andel, M., Gilbert, M., Robinson, T., & Robinson, A. (2018). How do you find the green sheep? A critical review of the use of remotely sensed imagery to detect and count animals. *Methods in Ecology and Evolution*, *9*(4), 881–892. https://doi.org/10.1111/2041-210X.12973

Horning, N. (2008). Remote Sensing. In S. E. Jørgensen & B. D. Fath (Eds.), *Encyclopedia of Ecology* (pp. 2986–2994). https://doi.org/10.1016/B978-008045405-4.00237-8

Hu, F., Xia, G.-S., Hu, J., & Zhang, L. (2015). Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. *Remote Sensing*, *7*(11). https://doi.org/10.3390/rs71114680

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794.

Jean, N., Wang, S., Samar, A., Azzari, G., Lobell, D., & Ermon, S. (2018). Tile2Vec: Unsupervised representation learning for spatially distributed data. *ArXiv Preprint ArXiv:1805.02855*.

Kostianoy, A., Litovchenko, K., Lavrova, O., Mityagina, M., Bocharova, T., Lebedev, S., … Pichuzhkina, O. (2006). Operational satellite monitoring of oil spill pollution in the Southeastern Baltic Sea: 1.5 Years experience. *2006 IEEE US/EU Baltic International Symposium*, 1–11. https://doi.org/10.1109/BALTIC.2006.7266136

Krizhevsky, A. (2009). *Learning Multiple Layers of Features from Tiny Images*. 60.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25* (pp.

1097–1105). Retrieved from http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

Längkvist, M., Kiselev, A., Alirezaie, M., & Loutfi, A. (2016). Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks. *Remote Sensing*, *8*(4). https://doi.org/10.3390/rs8040329

LeCun, Y., Fu Jie Huang, & Bottou, L. (2004). Learning methods for generic object recognition with invariance to pose and lighting. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, *2*, 97–104. https://doi.org/10.1109/CVPR.2004.1315150

Leichtle, T., Geiß, C., Wurm, M., Lakes, T., & Taubenböck, H. (2017). Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment. *International Journal of Applied Earth Observation and Geoinformation*, *54*, 15–27. https://doi.org/10.1016/j.jag.2016.08.010

Lein, J. K. (2009). Implementing remote sensing strategies to support environmental compliance assessment: A neural network application. *Environmental Science & Policy*, *12*(7), 948–958. https://doi.org/10.1016/j.envsci.2009.08.001

Li Fei-Fei, Fergus, R., & Perona, P. (2004). Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories. *2004 Conference on Computer Vision and Pattern Recognition Workshop*, 178–178. https://doi.org/10.1109/CVPR.2004.383

Li, T., Shen, H., Yuan, Q., Zhang, X., & Zhang, L. (2017). Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach. *Geophysical Research Letters*, *44*(23), 11,985-11,993.

https://doi.org/10.1002/2017GL075710

Liu, Z., Yuan, L., Weng, L., & Yang, Y. (2017). A High Resolution Optical Satellite Image Dataset for Ship Recognition and Some New Baselines. *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM,* 324–331. https://doi.org/10.5220/0006120603240331

Löffler, E., & Margules, C. (1980). Wombats detected from space. *Remote Sensing of Environment*, *9*(1), 47–56. https://doi.org/10.1016/0034-4257(80)90046-2

Lu, X., Zheng, X., & Yuan, Y. (2017). Remote Sensing Scene Classification by Unsupervised Representation Learning. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(9), 5148–5157. https://doi.org/10.1109/TGRS.2017.2702596

Matsuyama, T. (1989). Expert Systems for Image Processing: Knowledge-Based Composition of Image Analysis Processes. *Computer Vision, Graphics, and Image Processing*, *48*, 22–49.

Matsuyama, T., & Hwang, V. S.-S. (1990). *SIGMA: A knowledge-based aerial image understanding system*. Springer Science & Business Media.

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Mootz, J. (2017, November). *2019 NAIP*. Retrieved from https://www.fgdc.gov/organization/coordination-group/meeting-minutes/2017/november/naip-2019-cg-20171114.pdf

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(3), 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001

Mundhenk, T. N., Konjevod, G., Sakla, W. A., & Boakye, K. (2016). A Large Contextual Dataset for Classification, Detection and Counting of Cars with Deep Learning. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer Vision – ECCV 2016* (pp. 785–800). Springer International Publishing.

Nicole, W. (2013). CAFOs and environmental justice: The case of North Carolina. *Environmental Health Perspectives*, *121*(6), a182.

Onoda, M., & Young, O. R. (Eds.). (2017). *Satellite Earth Observations and Their Impact on Society and Policy*. https://doi.org/10.1007/978-981-10-3713-9

Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1717–1724. https://doi.org/10.1109/CVPR.2014.222

Purdy, R. (2010). Using Earth Observation Technologies for Better Regulatory Compliance and Enforcement of Environmental Laws. *Journal of Environmental Law*, *22*(1), 59–87. https://doi.org/10.1093/jel/eqp027

Purdy, Ray, Harris, R., Carver, J., & Slater, D. (2017). *Smarter Regulation of Waste in Europe*.

Rezaee, M., Mahdianpari, M., Zhang, Y., & Salehi, B. (2018). Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(9), 3030–3039. https://doi.org/10.1109/JSTARS.2018.2846178

Richards, J. A. (2013). *Remote sensing digital image analysis* (Vol. 5). Springer.

Rodgers, S., & Haines, J. (2005). *Detecting and Mitigating the Environmental Impact of Fecal Pathogens Originating from Confined Animal Feeding Operations*. Environmental Protection Agency.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach*. Malaysia; Pearson Education Limited,.

Salomon, M., Markus, T., & Dross, M. (2014). Masterstroke or paper tiger – The reform of the EU′s Common Fisheries Policy. *Marine Policy*, *47*, 76–84. https://doi.org/10.1016/j.marpol.2014.02.001

Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, *61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Skole, D., & Tucker, C. (1993). Tropical Deforestation and Habitat Fragmentation in the Amazon: Satellite Data from 1978 to 1988. *Science*, *260*(5116), 1905–1910. https://doi.org/10.1126/science.260.5116.1905

Slonecker, E. T., Jennings, D. B., & Garofalo, D. (2001). Remote sensing of impervious surfaces: A review. *Remote Sensing Reviews*, *20*(3), 227–255. https://doi.org/10.1080/02757250109532436

Sullivan, D. M., & Krupnick, A. (2018). Using Satellite Data to Fill the Gaps in the US Air Pollution Monitoring Network. *Resources for the Future Working Paper*, 18–21.

Syrris, V., Hasenohr, P., Delipetrev, B., Kotsev, A., Kempeneers, P., & Soille, P. (2019). Evaluation of the Potential of Convolutional Neural Networks and Random Forests for Multi-Class Segmentation of Sentinel-2 Imagery. *Remote Sensing*, *11*, 907. https://doi.org/10.3390/rs11080907

Union of Concerned Scientists. (2019). *UCS Satellite Database*.

Van Etten, A. (2018). You only look twice: Rapid multi-scale object detection in satellite imagery. *ArXiv Preprint ArXiv:1805.09512*.

Van Etten, A. (2019). Satellite Imagery Multiscale Rapid Detection with Windowed Networks. *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 735–743.

IEEE.

Van Etten, A., Lindenbaum, D., & Bacastow, T. M. (2018). Spacenet: A remote sensing dataset and challenge series. *ArXiv Preprint ArXiv:1807.01232*.

Wang, Q., Zhang, J., Hu, X., & Wang, Y. (2016). Automatic Detection and Classification of Oil Tanks in Optical Satellite Images Based on Convolutional Neural Network. In A. Mansouri, F. Nouboud, A. Chalifour, D. Mammass, J. Meunier, & A. Elmoataz (Eds.), *Image and Signal Processing* (pp. 304–313). Springer International Publishing.

Witte, J. C., Schoeberl, M. R., Douglass, A. R., Gleason, J. F., Krotkov, N. A., Gille, J. C., … Livesey, N. (2009). Satellite observations of changes in air quality during the 2008 Beijing Olympics and Paralympics. *Geophysical Research Letters*, *36*(17). https://doi.org/10.1029/2009GL039236

Xie, M., Jean, N., Burke, M., Lobell, D., & Ermon, S. (2016). Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping. *Thirtieth AAAI Conference on Artificial Intelligence*. Presented at the Thirtieth AAAI Conference on Artificial Intelligence. Retrieved from https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12196

Yifei Xue, Tiejun Wang, & Andrew K. Skidmore. (2017). Automatic Counting of Large Mammals from Very High Resolution Panchromatic Satellite Imagery. *Remote Sensing*, *9*, 878–894.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. *CVPR*.