

OPPORTUNITIES FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE IN NATIONAL MAPPING AGENCIES: ENHANCING ORDNANCE SURVEY WORKFLOW

J. Murray^{12*}, I. Sargent¹³, D. Holland¹, A. Gardiner¹, K. Dionysopoulou¹, S. Coupland¹, J. Hare³, C. Zhang², P.M. Atkinson²³⁴

¹Ordnance Survey, Explorer House, Adanac Drive, Nursling, Southampton, SO16 0AS, UK (Jon.Murray, Isabel.Sargent, David.Holland, Andy.Gardiner, Kyriaki.Dionysopoulou, Steven.Coupland)@os.uk

²Lancaster University, Bailrigg, Lancaster, LA1 4YQ, UK (j.murray3, c.zhang9, pma)@lancaster.ac.uk

³University of Southampton, University Road, Southampton, SO17 1BJ, UK (is, jsh2, pma)@ecs.soton.ac.uk

⁴Chinese Academy of Sciences, 11A Datun Road, Beijing 100101, China

KEY WORDS: Ordnance Survey, Artificial Intelligence, Machine Learning, Deep Neural Networks, National Mapping Agency.

ABSTRACT:

National Mapping agencies (NMA) are frequently tasked with providing highly accurate geospatial data for a range of customers. Traditionally, this challenge has been met by combining the collection of remote sensing data with extensive field work, and the manual interpretation and processing of the combined data. Consequently, this task is a significant logistical undertaking which benefits the production of high quality output, but which is extremely expensive to deliver. Therefore, novel approaches that can automate feature extraction and classification from remotely sensed data, are of great potential interest to NMAs across the entire sector. Using research undertaken at Great Britain's NMA; Ordnance Survey (OS) as an example, this paper provides an overview of the recent advances at an NMA in the use of artificial intelligence (AI), including machine learning (ML) and deep learning (DL) based applications. Examples of these approaches are in automating the process of feature extraction and classification from remotely sensed aerial imagery. In addition, recent OS research in applying deep (convolutional) neural network architectures to image classification are also described. This overview is intended to be useful to other NMAs who may be considering the adoption of similar approaches within their workflows.

1. INTRODUCTION

National mapping agencies (NMA) of the world are typically tasked with producing geospatial data and topographic maps of their respective countries. Due to the enormity of the task, which is coupled with the requirement to produce high quality data, many NMAs utilise a combination of remote sensing data capture with field survey activities to extract an extensive range of real-world features and characteristics. Remote sensing activities are focused predominantly on the acquisition of highly detailed aerial imagery. For example, this can be imagery at a pixel spatial resolution of 25cm resulting in several thousand rows and columns per image (Sargent et al., 2019). Resultingly, these images contain greater levels of detail and information than it is possible for NMAs to extract and make available to their customers using traditional, part-manual processing methods (Holland & Marshall, 2004, Cygan, 2019, Sargent et al., 2019).

Like many other NMAs, Great Britain's Ordnance Survey (OS) is embracing research opportunities to move towards full automation, while maintaining the focus on enabling the delivery of authoritative geospatial data and topographic mapping (Cygan, 2019). Typical for NMAs which also provide geospatial infrastructure data for their respective countries, OS provides mapping services for UK government, businesses and individual consumers, and products and services that rely upon a data capture and processing workflow that costs tens of millions of pounds to operate. Until recently, the capture and maintenance of geospatial data at OS was predominantly a manual process. However, OS research interests have turned towards optimising the information flow from the source data and identifying the

potential of artificial intelligence (AI) automation in the workflow to enhance the product offerings to customers. Using OS as an example NMA, this paper briefly describes the development of past, present and future AI projects that will achieve positive impacts on a typical workflow within an NMA.

2. ISSUES OF LARGE-SCALE GEOSPATIAL DATA COLLECTION FOR AN NMA

Creating geospatial products for NMA customers is a massive undertaking. For example, OS captures aerial imagery covering approximately 80-90,000 km² of the United Kingdom annually. This results in a weekly workflow of approximately 100,000 (change) updates, iterated over 650,000,000 features (Ordnance Survey, 2020). This task therefore, presents the significant problem of how to manage the assessment and correction of such a large number of features. In undertaking attempts to tackle this problem, OS undertook a sustained period of research into the automation of the change detection process (Holland, Gladstone et al., 2012). Subsequently, OS developed a rule-based automation process utilising eCognition (eCognition Essentials, Trimble, 2015) to increase efficiencies within the change detection workflow. These rule-based improvements detected change to a 92% correctness value. The change detection system can also create false positives but these are quickly discarded in a subsequent manual process, leading to significant processing time savings: ~50% time reduction, when compared to the equivalent manual process (Holland, Gladstone et al., 2012). In addition, other robust, automated methods of extracting feature attribution, such as building heights, have been added successfully to production pipelines and have also expanded

* Corresponding Author

OS's product portfolio for customers (Sargent et al., 2015). These improvements highlight the potential benefits of more efficient processing strategies for NMAs, where service improvements can be made through time-based efficiency savings or by facilitating increased amounts of geospatial data that can be pushed through the workflow. This has direct benefits for both the business and the customer, and has further driven OS towards increasing the use of machine learning (ML) and AI within the operational workflow.

3. EARLY EXPLORATION OF AI CAPABILITIES AT OS

OS research into the AI field began in 2015, where research focussed on understanding the capabilities of such techniques and investigating how AI could be used to extract additional information from previously acquired aerial image data, such as manually captured topographic data during photogrammetric surveys. It was understood that AI would enable greater understanding of our world, through the identification, disentangling and explanation of the many influencing factors found in environmental remotely sensed data (Bengio et al., 2013). Advances in machine learning had predominantly been achieved using limited, specialised datasets (e.g. ImageNet, Deng et al., 2009). Therefore, OS explored how representation learning could be applied to NMA datasets (Sargent, 2019). Early research applying machine learning techniques focused on unsupervised approaches to categorising roof shapes from digital surface models (Sargent et al., 2015). Some simple roof shapes and building shapes were identified, as well as artefacts such as overhanging vegetation which are useful to identify to reduce the instances of label misclassification. This initial research provided a series of opportunities that informed the AI research direction of the OS from this point forward.

4. ENHANCING FEATURE RECOGNITION WITH AI

To further the potential of AI use for feature recognition, a prototype two-phase deep learning (DL) algorithm, trained using aerial imagery and OS topographical data, was developed at OS in collaboration with University of Southampton and Lancaster University. The focus was to enable OS to extract key attribute information from aerial imagery. The first phase creates a general-purpose model for pre-processing imagery such that the outputs can be subsequently used as inputs to model the required product, for example, building attributes (Sargent et al., 2019). This method, TopoNet (Figure 1), is a deep neural network which identifies characteristic, repeated patterns from large scale aerial imagery (Sargent et al., 2019). Similar to other classification methods that utilise computer vision and DL (Branson et al., 2018, Griffiths & Boehm 2019), TopoNet utilises a deep convolutional neural network (DCNN) to act as a feature extractor, where multiple layers of convolutional filters are learned using back-propagation within a Keras framework (Chollet, 2018). Using this approach in the processing of aerial imagery permits the fluid manipulation of pre-constructed network architecture. During initial experimentation, the 8-layer AlexNet (Krizhevsky 2012) was tested, and subsequently when this architecture was exploited to its maximum potential, ResNet-50 was utilised as the feature extractor. As the name suggests, ResNet-50 provides a 50-layer network which also has the capacity to utilise the 'skip connection' method. This resolves potential gradient vanishing problems by reducing data extinction, instead of simply stacking convolution layers one after another (He et al., 2015).

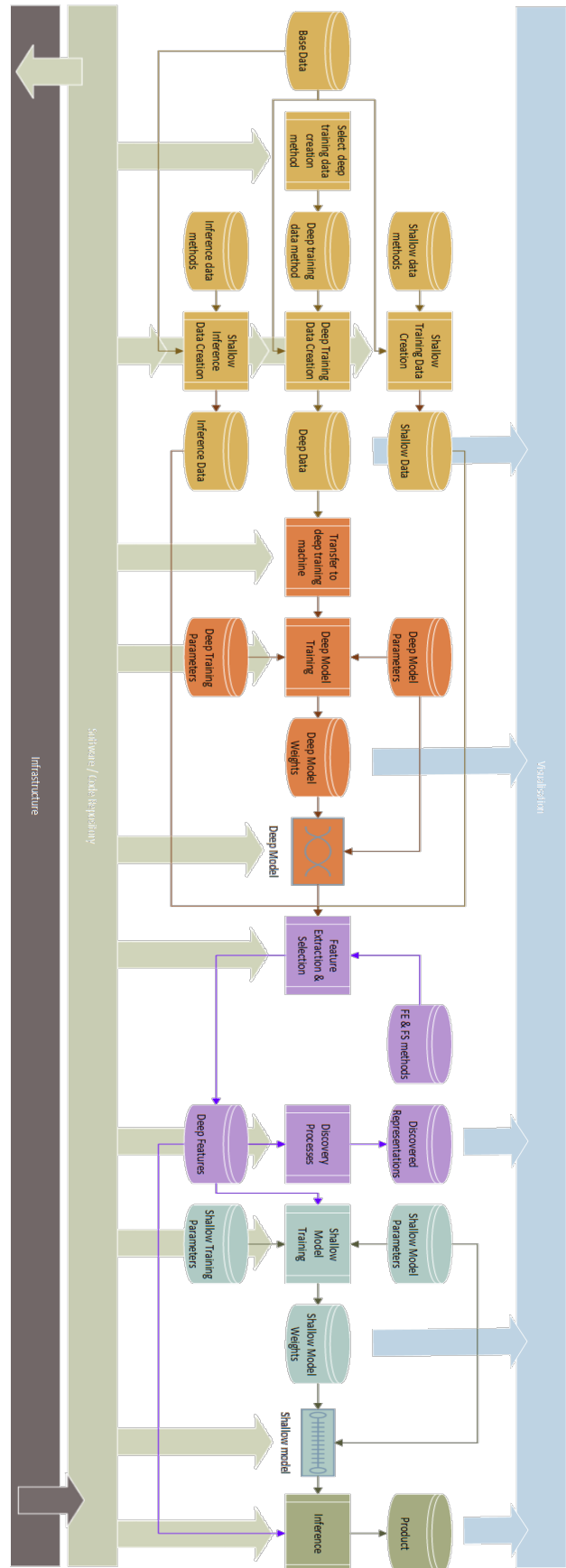


Figure 1. A model of the core components of the TopoNet deep learning method used at Ordnance Survey, UK.

The secondary phases of TopoNet use the techniques of ‘inference’ and ‘discovery’ (Figure 1), where inference utilises shallow machine learning approaches to respond to bespoke requests from customers and discovery involves investigation of ways to better understand the landscape (Sargent et al., 2017). It should be noted that *inference* is based on a technique commonly referred to in the literature as transfer learning (Weiss et al., 2016). OS efforts focused mainly on reducing the computational time needed to train weights for complex architectures for bespoke customer requests. This was achieved by taking advantage of weights trained for one task and utilising them to either a) initialize the weights for another network trained on a similar task, or b) create a simpler, and faster to compute, surrogate model that encapsulates important information learnt on the first task. In the latter case, the encapsulated information can subsequently be fed into a shallower network specific to the customer’s needs.

Transfer learning has a regularising effect on the model weights. This implies that a model obtained with weights that were initialized with weights that were pre-trained on a similar task, will overfit less to the data and therefore achieve a better performance on unseen data. Finally, transfer learning can be particularly beneficial in cases where labelled data are scarce, for example, when there is a customer need but the customer does not have a lot of labelled data or when data labelling has a prohibitive timeframe or cost associated to it. Transfer learning can yield extraordinary results when the content of the datasets, on which the new model is trained, is similar. The more dissimilar the content is, the more reference to earlier layers of the original network must be made in order to obtain a descriptive shallow model. Therefore, it is important to investigate and discover what the network has learnt in each one of the layers and how relevant the representations are to the task at hand. By performing both discovery as well as inference, it is hoped that deeper and more meaningful landscape understanding can be obtained to address longer-term customer requirements (Sargent et al., 2017).

Because TopoNet is a deep network, trained using a very large dataset (~1.2M examples), training is computationally intensive and can take many hours, or potentially weeks, depending on the complexity of the task and the computer infrastructure available. The initial network weights were based on ImageNet, but later versions of TopoNet used increasingly refined aerial image datasets. Features extracted from each version of TopoNet could subsequently be compared for use as input for the inference phase, such as classification accuracies when inferring building features (Figure 2). Changes to how the example image patches were labelled for TopoNet V2 and V3 produced improvements to the network that resulted in notably increased inference accuracy (Figure 2 a.) and it was also found that using features extracted from non-final layers produced the best results (Figure 2 b. and c.).

Preliminary experimentation applied a range of techniques to *discover* what was learnt by the network and produced intriguing results. For example, through examining the image clips within the input data set that were most responsible for the firing of a particular node in the network, provides insight into which network representations are learnt. Figure 3 highlights a selection of image clips and visualises how the nodes can respond to characteristics in the image set. Additional research is underway to establish how meaningful these different representations are.

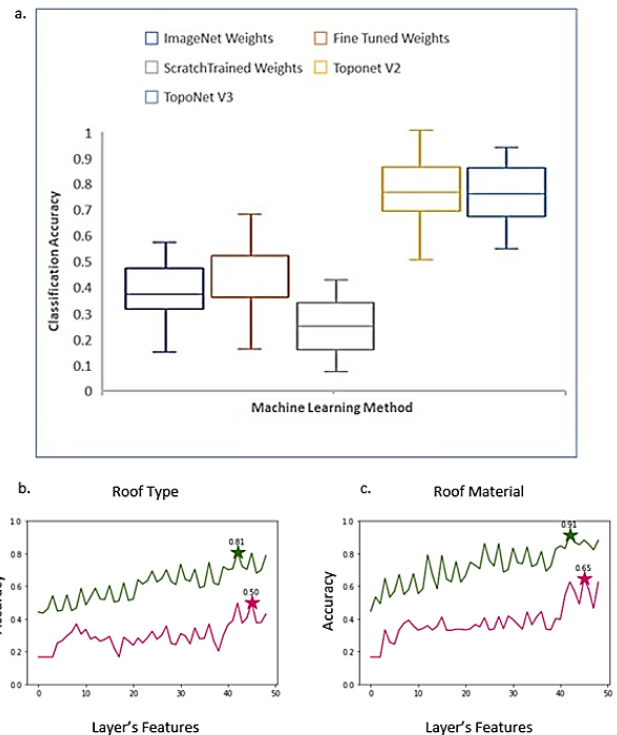


Figure 2. Comparison of classification accuracy from features extracted by the deep neural network, TopoNet. a. overall classification accuracy of recent versions of TopoNet (orange and blue) when compared against other deep learning networks (ImageNet, and two early experiments at training deep networks with aerial imagery). b. accuracy for inference of roof types, and c. accuracy for inference of roof material, comparing both ImageNet and TopoNet methods. Legend (b & c): Pink line: ImageNet features, Green line: TopoNet features

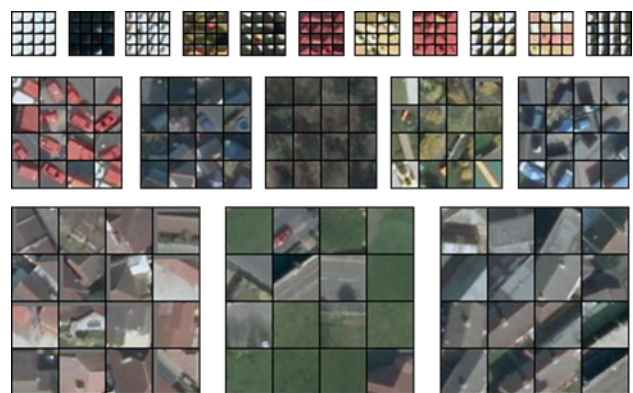


Figure 3. Examples of node responses within a CNN trained with aerial imagery. Each of the 19 squares contains the top 16 image clips that caused the highest activation of a given node within the network. *Top Row*: The 11 smallest squares are from one of the earliest activation layers in the network and demonstrate how early layers respond to edge features within the imagery. *Middle Row*: These 5 squares from a middle layer within the network, and apparently respond to specific colours. *Bottom Row*: This row shows how nodes may demonstrate more ‘semantic’ responses such as roofs and fields.

5. THE FUTURE OF AI IN NMAS

OS has undertaken several exploratory experiments assessing whether deep networks can learn meaningful representations of the landscape. In these experiments, broad land use and land cover categories (e.g. industrial, residential, water, grassland) separate well, and OS is currently interrogating this result by defining the result's semantic underpinning (i.e. is this due to object colour, texture?, etc.). AI can identify repeated patterns in the landscape and this is seen as a key research area for the future recognition of land use categories. OS has initiated a collaboration with the National Physical Laboratories and Science and Technology Facilities Council, aiming to gain greater understanding of landscape representations learned by deep networks using remotely sensed imagery. It is envisaged that by using the TopoNet approach, OS and other NMAs will learn new ways to discover landscape features (Sargent et al., 2019). For example OS is contributing to novel projects for exploiting the maximum potential from geospatial data, by aiming to understand how to uncover extremely obscured landscape features, OS has sponsored research into applying deep learning within archaeology (Kramer, Hare, et al., 2017).

Furthermore, a new concept of utilising object-based convolutional neural networks (OCNN) has been developed in partnership with Lancaster University. OCNN, combines CNN and object-based image analysis (OBIA), and demonstrates the potential to classify complex land uses through deep feature representations, while maintaining fine spatial detail through regional partition and boundary delineation. This provides new perspectives on object description and feature characterisation, where both within-object information and between-object information are jointly learned (Zhang et al., 2018). OCNN approach has achieved classification accuracies at ~90%, which is a significant increase over other established classification methods such as Markov Random Field (MRF), OBIA-SVM and pixel-based CNN methods (Zhang et al., 2018) (Table 1). OCNN is also used as a preparatory step within Joint Deep Learning (JDL). The JDL approach utilises Markov iteration to update between land cover (LC) and land use (LU) classifications (figure 4).

Class	MRF	OBIA-SVM	Pixel-wise CNN	OCNN 48*	OCNN 128	OCNN 128+48*
Commercial	70.09	72.87	73.26	76.4	81.13	82.46
Highway	77.23	78.04	76.12	78.17	74.35	79.69
Industrial	67.28	69.01	71.23	78.24	83.87	84.75
High-density	81.52	80.59	80.05	81.75	85.35	86.43
Residential Medium density	82.54	84.42	85.27	87.28	90.34	90.59
Residential Parks and Recreation	91.05	93.14	92.34	92.59	96.41	97.09
Parking	80.09	83.17	84.76	86.02	85.59	88.83
Railway	88.07	90.65	86.57	89.51	87.28	91.92
Redeveloped Area	89.13	90.02	89.26	89.71	94.57	94.69
Harbour and Sea	97.39	98.43	98.54	98.62	98.75	98.95
Overall Accuracy (OA)	78.67%	79.54%	81.62%	84.23%	87.31%	89.52%
Kappa Coefficient (k)	0.76	0.78	0.8	0.82	0.86	0.88

Table 1. Classification accuracy comparison between MRF, OBIA-SVM, Pixel-wise CNN, OCNN48*, OCNN128, and

OCNN128+48*. Overall accuracy (OA) and Kappa coefficient (k). 'Bold' highlights the highest accuracy per method

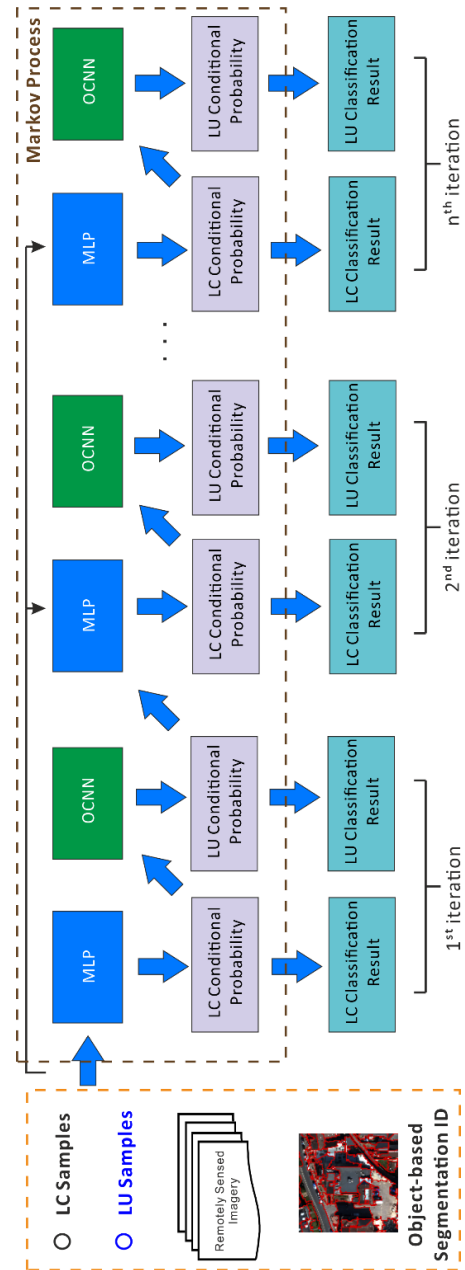


Figure 4. The joint deep learning (JDL) to classify land cover and land use in a single unified framework.

The JDL framework represents a new paradigm in remote sensing classification, where the previously separate goals of LC (state i.e. what is there?) and LU (function i.e. what is happening there?) are brought together in a single unifying framework. This allows both the pixel-based low-order and higher-order representations to interact and update iteratively, which complements the mutual refinement of both the LC and LU classifications. Initial research shows that this method results in further increases in classification accuracy and has the potential to enhance the generalised processing workflows of NMAs at a range of data levels or topographic scales (Zhang et al., 2019). It is further suggested that a key advantage of the JDL method is

the move away from a 2-part classification process, specifically the utilisation of both LC and LU data (Zhang et al., 2019), which could achieve higher classification accuracies and improved workflows for NMAs.

6. CONCLUDING REMARKS

It is understood that to maximise the potential for AI use in NMAs, robust systems of network training and interrogation need to be developed to understand where AI supported discoveries are meaningful and to what end these could be applied in an operational sense (Sargent et al., 2019). It is understood that there is value in utilising AI and ML in the operational workflows of NMAs, and this can lead to significant business efficiencies, greater product consistency and an enhanced series of products available for the customer.

Throughout their exploration of developing AI use within the workflow, OS has taken the opportunity to learn the unique requirements of a suitable computational infrastructure which will enable the handling of the AI data flow and permits the robust development of workflow automation. These initial forays into AI automation have provided great insights into the potential application of AI for NMAs and, therefore, OS plans to further maximise the benefits of developing AI capabilities in the near future, which it is envisaged will lead to greater uptake and application of AI use within the NMA sector.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their insightful comments in reshaping this paper. The research associated with National Physical Laboratories and Science and Technology Facilities Council is funded by Innovate UK under the A4I (Analysis for Innovators programme) project number: 36335. This research was jointly funded by the Engineering and Physical Sciences Research Council; Impact Acceleration Account and Ordnance Survey. The lead author would also like to acknowledge the assistance of Miss B. Heap, Lancaster University, UK, for help given during the preparation of this manuscript.

REFERENCES

Bengio, Y., Courville, A., Vincent, P., 2013. Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Volume 35, no 8. 1798-1828. doi: 10.1109/TPAMI.2013.50.

Branson, S., Wegner, J.D., Hall, D., Lang, N., Schindler, K., Perona, P., 2018. From Google Maps to a fine-grained catalog of street trees. *ISPRS Journal of Photogrammetry and Remote Sensing*. Volume 135, January 2018, 13-30. doi.org/10.1016/j.isprsjprs.2017.11.008

Chollet, F., 2018. Keras Documentation. <https://keras.io>. (4 May 2020)

Cygan, M., 2019. ESRI Blog. <https://www.esri.com/about/newsroom/blog/great-britains-national-mapping-agency-embraces-automation/> (4 May 2020)

Deng, J., Dong, W., Socher, R., Li, L., Li, K., Fei-Fei, L., 2009. ImageNet: A Large-Scale Hierarchical Image Database. *IEEE Computer Vision and Pattern Recognition (CVPR)*. 248-255. doi: 10.1109/CVPR.2009.5206848

Griffiths, D., Boehm, J., 2019. Improving public data for building segmentation from Convolutional Neural Networks (CNNs) for fused airborne lidar and image data using active contours. *ISPRS Journal of Photogrammetry and Remote Sensing*. Volume 154, August 2019, 70-83. doi.org/10.1016/j.isprsjprs.2019.05.013

Holland, D., Gladstone, C., Sargent, I., Horgan, J., Gardiner, A., Freeman, M., 2012. Automating the Photogrammetric Workflow in a National Mapping Agency. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences I-4*: 83-88. doi.org/10.5194/isprsannals-I-4-83-2012

Holland, D. and Marshall, P., 2004. Updating maps in a well-mapped country using high resolution satellite imagery. *Int. Arch. Photogramm., Remote Sens. Spatial Inform. Sci.* 35.

Kaiming, H., Zhang, X., Ren, S., and Sun, J., 2015. Deep Residual Learning for Image Recognition. arXiv:1512.03385. (4 May 2020).

Kramer, I. C., Hare, J., Prugel-Bennett, A. and Sargent, I., 2017. Automated detection of archaeology in the New Forest using deep learning with remote sensor data. *New Forest Knowledge Conference 2017: New Forest Historical Research and Archaeology: Who's doing it?* October 2017, Lyndhurst, UK.

Krizhevsky, A., Sutskever, I., Hinton, G. E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* 1097-1105. <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks> (4 May 2020).

Ordnance Survey. 2020. Ordnance Survey History: 1783-1791. <https://www.ordnancesurvey.co.uk/about/history>. (3 February 2020).

Sargent, I., Doidge, C., Wilson, O., Rainbow, J. 2019. Ordnance Survey's Approach to Machine Learning for Inference and Discovery using Imagery. *Geomatics World*. Hertfordshire, UK., Geomares Publishing LTD. 27, 20-23.

Sargent, I., Hare, J., Young, D., Wilson, O., Doidge, C., Holland, D., Atkinson, P. 2017. Inference and Discovery in Remote Sensing Data with Features Extracted Using Deep Networks. *SGAI 2017*, Cambridge, England, SGAI.

Weiss, K., Khoshgoftaar, T.M., Wang, D. (2016) A Survey of Transfer Learning. *Journal of Big Data*, 3, 9. doi.org/10.1186/s40537-016-0043-6

Wignall, M. 2019. "Raising the Roof with Artificial Intelligence. <https://www.ordnancesurvey.co.uk/business-government/innovation/happens/articles-ehs/raising-the-roof-with-artificial-intelligence-ne>. (5th May 2020)

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P.M. 2018. An object-based convolutional neural network (OCNN) for urban land use classification *Remote Sensing of Environment*, 216, 57-70.

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P.M., 2019. Joint Deep Learning for land cover and land use classification. *Remote Sensing of Environment*, 221. 173-187.