Heriot-Watt University
Research Gateway

Aerial Image Analysis Using Deep Learning for Electrical Overhead Line Network Asset Management

Citation for published version:
Odo, A, McKenna, S, Flynn, D & Vorstius, JB 2021, 'Aerial Image Analysis Using Deep Learning for Electrical Overhead Line Network Asset Management', *IEEE Access*, vol. 9, pp. 146281-146295. https://doi.org/10.1109/ACCESS.2021.3123158

Digital Object Identifier (DOI):
10.1109/ACCESS.2021.3123158

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Publisher's PDF, also known as Version of record

Published In:
IEEE Access

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 25. Mar. 2025

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 31, 2021, accepted October 17, 2021, date of publication October 26, 2021, date of current version November 3, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3123158

Aerial Image Analysis Using Deep Learning for Electrical Overhead Line Network Asset Management

ANICETUS ODO[1], STEPHEN MCKENNA[1], DAVID FLYNN[2], (Member, IEEE), AND JAN BERND VORSTIUS[1]

[1]School of Science and Engineering, University of Dundee, Dundee DD1 4HN, U.K.
[2]Smart Systems Group, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K.

Corresponding author: Stephen McKenna (s.j.z.mckenna@dundee.ac.uk)

The work of Anicetus Odo was supported by the Tertiary Education Trust Fund (TETFund) of Nigeria. The work of David Flynn was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) National Centre for Energy System Integration Project under Grant EP/P001173/1.

ABSTRACT Electricity networks are critical infrastructure, delivering vital energy services. Due to the significant number, variety and distribution of electrical network overhead line assets, energy network operators spend millions annually on inspection and maintenance programmes. Currently, inspection involves acquiring and manually analysing aerial images. This is labour intensive and subjective. Along with costs associated with helicopter or drone operations, data analysis represents a significant financial burden to network operators. We propose an approach to automating assessment of the condition of electrical towers. Importantly, we train machine learning tower classifiers without using condition labels for individual components of interest. Instead, learning is supervised using only condition labels for towers in their entirety. This enables us to use a real-world industry dataset without needing costly additional human labelling of thousands of individual components. Our prototype first detects instances of components in multiple images of each tower, using Mask R-CNN or RetinaNet. It then predicts tower condition ratings using one of two approaches: (i) component instance classifiers trained using class labels transferred from towers to each of their detected component instances, or (ii) multiple instance learning classifiers based on bags of detected instances. Instance or bag class predictions are aggregated to obtain tower condition ratings. Evaluation used a dataset with representative tower images and associated condition ratings covering a range of component types, scenes, environmental conditions, and viewpoints. We report experiments investigating classification of towers based on the condition of their multiple insulator and U-bolt components. Insulators and their U-bolts were detected with average precision of 96.7 and 97.9, respectively. Tower classification achieved areas under ROC curves of 0.94 and 0.98 for insulator condition and U-bolt condition ratings, respectively. Thus we demonstrate that tower condition classifiers can be trained effectively without labelling the condition of individual components.

INDEX TERMS Pattern recognition, machine learning, image classification, asset management, power transmission.

I. INTRODUCTION

Transmission and distribution of electricity are critical energy services for communities globally, underpinning vital services such as telecommunications, water services, transport and education. Electrical network overhead line (OHL) assets are inspected regularly for failures or conditions that might lead to faults. This is done for safety and economic reasons, and because it is required by law. In the UK, for example, Section 24 of the Electricity Supply Regulation 1988 requires distribution network operators to take practical power-line inspection steps to avoid unplanned outages [1]. Regulators of the energy industry have developed a set of standards for assessing network assets. Electrical towers are inspected against parameters that encompass the condition of their various component parts such as insulators, conductors and

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wang.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

U-bolts, as well as factors such as bird nesting and accumulation of droppings, and encroachment of vegetation. Inspection of high voltage line assets is especially challenging, requiring extreme caution. The use of people on patrol for monitoring OHL corridor and tower components is now out of fashion because of the high risk and relative inefficiency in terms of response time. Efforts to improve inspection quality and safety, and reduce costs and risks of failure, focus on exploiting optical aerial and satellite imaging and airborne laser scanners [2]. Low-altitude aerial imaging is used for fine-grained inspection with use of helicopters commonplace in the industry. Eyre-Walker *et al.* [3] described advanced condition assessment and asset management techniques adopted by distribution network operators in the UK as key components in inspection and refurbishment of OHL assets. More recently, there has been increasing use of drone technology with a short-term, single-mission focus [4]. A drone can be flown close to an asset and take high-resolution images from different views. Alternatives to drones in terms of flexibility and cost are roll-on-wire and hybrid climbing-flying robots. Although there are proposals for the deployment of these alternatives for OHL asset inspection, they are yet to be adopted as business as usual (BaU) capabilities. Current inspection pipelines capture aerial images via helicopters and the acquired aerial images are manually analysed by experts for faults or precursor signatures of faults. This is a highly labour-intensive and subjective process, that is error prone and expensive.

In this paper, we propose an approach to automated tower condition assessment from aerial images based on deep learning. We describe prototype systems and report results evaluating performance on real-world industry data. Importantly, the machine learning methods used to predict tower condition ratings are trained without using condition rating labels for individual components of interest. Instead, learning is supervised using only condition labels for each tower as a whole. This approach enables us to use a real-world industry dataset without needing costly additional human annotation of thousands of individual components' conditions.

Experiments in this paper focus on assessing towers in terms of the condition of their insulator components and their U-bolt components. These components are used as exemplars to demonstrate our approach which we believe will also be applicable to various other parameters of interest in future work. Insulator faults can arise from cracking, chipping, shattering, and rust. U-bolts are used to attach suspension insulators to a tower and are susceptible to wear and rust.

We highlight the importance of monitoring and the role of automation with a review of relevant literature focused on the detection and analysis of OHL assets from aerial images (Section II). An important aspect of our study is the use of a real-world industry dataset with representative tower images and associated condition ratings covering a range of component types, scenes, environmental conditions and viewpoints. We provide an overview of this dataset in Section III.

In Section IV, we give a system overview of our proposed approach to inspection. This involves first detecting instances of the components of interest in multiple images of each tower using deep learning detector networks. The condition ratings for each tower are then predicted using deep classifiers. Two methodologies for tower classification are compared. The first uses component instance classifiers trained using class labels transferred from the tower to each of its detected component instances. The second uses multiple instance learning (MIL) to train classifiers based on bags of detected instances. Instance or bag class predictions are aggregated to obtain tower condition ratings.

The deep networks used for instance detection are described and evaluated in Section V. Specifically, we compare Mask R-CNN and RetinaNet methods for the detection of insulators and two categories of U-bolt. We know of no previous published study comparing these methods for OHL tower inspection. Furthermore, automated detection of U-bolts from aerial images has not been previously reported. We also report experiments investigating how detection performance is affected by partial occlusion.

Section VI describes the deep networks used for instance-based and MIL-based tower classification. This is, to the best of our knowledge, the first time multiple instance learning has been used for inspection of electrical OHL assets. We propose the use of MIL on sub-bags of images to enable learning with constrained resources. Whereas previous studies have identified individual instances of faulty insulators based on annotated faults [5], [6], our approach avoids the need for costly labelling of the fault condition of individual components. Tower classification results reported in Section VII show that high accuracy condition ratings can be obtained, and that tower level labels can be adequate for the task. Finally, we discuss the findings of this study (Section VIII) and draw conclusions to inform further research (Section IX).

II. MOTIVATION AND RELATED WORK

A. THE IMPORTANCE OF ELECTRICAL ASSET MONITORING

The economic and societal implications of interrupting electrical supply have been demonstrated throughout the world [7]. Accounts relating to failure of electrical networks include those of blackouts in North America [8] and Europe [9]. An assessment of the impact of a 2003 blackout in Italy estimated the damage to society to exceed 1.15 billion Euros [9]. The root causes of such power failures are diverse and include loss of energy generation, switching failures, IT failures, and failures of OHL network assets. Degradation of OHL networks can cause disastrous environmental events such as a 2019 California wildfire reported to have burned over 150,000 acres and killed 85 people [10]. Component failures and human errors were reported to be responsible for a third of blackouts world-wide between 2011 and 2019 [7].

OHL electrical networks are complex, varied and highly distributed networks of assets such as electrical cables, towers, insulators, pole mount transformers (PMT) and arrestors.

Distributed generation is increasing within electrical networks, and this creates more complex bidirectional energy flows for a system that was originally designed to provide a one-way centralised energy service [11]. To support the continuity of reliable, affordable and increasingly more sustainable electrical energy, network reinforcement may be required and this represents a significant cost to network operators and consumers. For example, modernisation of electrical networks in the United States is expected to incur a $3 trillion cost, which excludes generation asset investment [12]. To support the deferment of network investment, and to ensure continuity of reliable service within a modernised electrical network, asset management has a strategic role. Miguelanez-Martin and Flynn [13] highlight the transition from time-based condition monitoring to predictive maintenance within the energy sector. Matikainen *et al.* [2] also note that effective asset monitoring is integral to continuity of service.

B. PREVIOUS WORK ON AUTOMATED MONITORING

We previously reported a method to automatically classify tower configurations from aerial images [14]. In this paper, we focus on the specific challenges of equipment failure on towers, specifically of insulators and U-bolts. Some researchers have focused on environmental factors that may lead to component degradation and failure. This has included detection of activities of birds on OHL towers [15], [16], detection of icing [17]–[21], and measurement of vegetation encroachment [22]. Others have addressed the detection of components that have failed or that require immediate maintenance work. This has included detection of broken line strands [23]–[25] and measurement of sagging on line segments [26], [27]. However, insulators are the components that have been most studied in terms of automating detection and condition assessment on OHL towers [28]. Studies have focused on detecting, and in some cases segmenting, insulators in images, with emphasis on electrical failure modes such as shattered or exploded sheds [29]–[32], missing or cracked insulator caps, or flash-over hot-spots [33]–[37]. We detect insulators and analyse insulator mechanical failure modes at the tower level, classifying towers as healthy or unhealthy based on the condition of the detected insulators, rather than segmenting component sub-regions such as missing insulator caps. Whereas Jalil *et al.* [38] train a condition detector using rating labels for individual insulators, we focus on aggregated tower condition rating. Another component type highlighted as high-risk by distribution network operators is U-bolts. The automated detection and analysis of U-bolts has not been presented previously in literature.

Methods for analysing images of OHL assets are drawn from the computer vision and deep learning literature. Traditional image analysis has had a central role, relying heavily on engineered extraction of texture, shape [39], [40], colour features [35], [37], [41], [42]. Fusion of features such as grey-levels, colour names and histogram of oriented gradients (HOG) has been used to enhance insulator detection performance [43]. Yu *et al.* [44] highlighted the difficulty of segmenting glass insulators in low contrast images and combined texture and shape features for the task. Insulators are diverse in shape, colour and texture, and handcrafted methods may have difficulty generalising over this diversity [31].

Deep learning bypasses the requirement to manually define discriminating features. Several convolutional neural network (CNN) architectures have been used for detection of OHL assets. These can be categorised as one-stage and two-stage networks. Two-stage networks typically incorporate two CNNs: the first is used for generating region proposals and the second acts to localise and classify objects. Examples of two-stage networks are Faster R-CNN [45] and Mask R-CNN [46] which achieved state of the art results on object detection benchmarks. One-stage networks include single-shot multibox object detection, SSD [47], and YOLO [48]. Evaluations on computer vision benchmarking datasets show that although two-stage networks can have better precision and can detect small objects in images, one-stage models are faster. Gao *et al.* [30] proposed analysis of insulators using a VGG16-based CNN to suggest regions which were fed into Faster R-CNN for insulator detection. Faster R-CNN was also applied to the detection of faulty insulator caps [30], [31]. Whereas [30], [49] worked on porcelain and composite insulator types, Ling *et al.* [31] explored glass insulators. As pointed out in [50], it is common to find a tower with a combination of insulator strings. Li *et al.* [5] used a similar method to [31] for global detection and local segmentation of glass and ceramic insulators but added online hard example mining to deal with class imbalance between foreground and background. The mined examples were then forward propagated through a Faster R-CNN for insulator detection and a U-net for segmentation. Sampedro *et al.* [6] detected insulators using a network derived from a modified U-net, trained within a generative adversarial network framework with transfer learning. Nguyen *et al.* [51] highlighted speed as a major motivation for using one-stage networks. Miao *et al.* [49] fine-tuned an SSD previously trained on the COCO dataset [52] with some aerial images comprising of different insulators and backgrounds. The model was further fine-tuned using a more specific data set of porcelain and composite insulators and considering scenes with vegetation, roof-tops, etc. YOLO has also been used for insulator detection [50], [53]. Han *et al.* [54] used a ResNet50 to extract insulator features for YOLO, and considered different input scales. Chen *et al.* [53] enhanced blurred images using a Super-Resolution Convolutional Neural Network (SRCNN). RetinaNet was introduced to bridge the gap between two-stage and one-stage models arising from class imbalance between foreground and background [55]. This method increases accuracy without compromising on speed but its application for the detection of OHL tower components has been limited to detection of pins from pre-processed tower [56]. Insulator detection using deep learning forms one part of our proposed system. However, our focus is on

assessing condition ratings at the tower level by aggregating information from multiple instances of insulators and other components without the need to label the conditions of individual components in the image data.

The amount and quality of data are important factors for deep learning methods. Unfortunately, the process of collecting and annotating aerial images is costly and there are several levels of restrictions especially with regards to protecting company assets and data protection. Authors use different proprietary data sets, which are often insufficient in size. Data augmentation has been used to artificially increase the size of small data sets [30], [31], [36], [38], [49], [50], [54] as well as to address imbalance between normal and abnormal targets, e.g., 60 defective insulators photographed from a UAV were used to generate a further 996 images using data augmentation in [36]. Images have also been sourced from the web [50]. Zhou *et al.* [57] used a large number of images for insulator and vibration hammer detection but how these images were distributed across different towers was not presented. It is important to know how components are distributed across tower instances because components on a given tower will have been exposed to similar environmental conditions and would therefore degrade at the same rate. Fusion of data from multiple sensors has also been used, e.g., visual and infrared image data [58]. Our experiments use a real-world electricity network inspection dataset comprising of visual images that are representative of the diverse component types and failure modes encountered in real-world inspection scenarios.

III. DATASET

The inspection database provided for this research is based on high-resolution still images taken during helicopter inspection surveys of electricity overhead lines [3]. Image acquisition involved a camera operator and an observer both of whom were engineers with good understanding of the information required for the inspection task. Photographs were taken using a digital camera fitted with an optically stabilised telephoto lens. To capture sufficient detail, multiple images of left and right circuits were taken with focus on condition parameters such as the earth-wire, conductors, tower paintwork, tower footing and vegetation encroachment. The observer was responsible for highlighting flash points that required more attention. The flights were carried out during spring and summer months between April and October when weather and lighting conditions were likely to be better. On average, 53 images were taken of each tower.

Three datasets, T6775, T7638, and T7799, were provided, each consisting of tower condition information and images from a different network survey. The surveys were conducted in 2011, 2014 and 2015, respectively. Towers inspected in T6775 and T7638 were from high voltage lines. Both low and medium voltage towers were inspected in T7799. The data sets comprises of 6,753 unique towers (i.e., 4,823, 993 and 937 towers in T6775, T7638 and T7799, respectively). A total of 233 towers were assessed twice across the data sets. For example, 219 towers surveyed in T6775 were scheduled again in T7638. Also, 14 towers inspected in T6775 are contained in surveys carried out as T7799. These towers were surveyed at different times. Hence the number of separate tower inspections is 7,219. Multiple condition ratings were assigned by experts to each tower based on inspection of its images. Each condition parameter was assessed based on guidelines agreed between the network operator and the service provider. Condition information was recorded in spreadsheets in which hyperlinks to associated tower images were provided.

TABLE 1. Numbers of images in the database, categorised by tower viewpoint.

Image category	Number of images	Percentage
Insulator footing	59,505	15.40
Insulator top	59,306	15.35
Insulator middle	58,805	15.22
Middle	46,133	11.94
Footing	39,281	10.17
Top	38,152	9.87
Anti-climb guard	31,942	8.27
Earth wire	30,735	7.95
Whole tower	*14,943*	*3.87*
Signage	*4,093*	*1.06*
Cable platform	*2,553*	*0.66*
Substation	*935*	*0.24*
Spacer	*18*	*0.00*
Total	386,401	100.00
Total selected	363,859	94.17

A. IMAGE DATA

Each tower has multiple images which vary in their viewpoint and thus capture different aspects and parts of the tower. A viewpoint category label was provided with each image. Table 1 lists these along with the total number of images in each category. *Whole tower* images were excluded because they were typically long-shots meaning that specific components were often not well resolved. We also excluded images specifically targeting *signage*, *cable platforms*, *substations* and *spacers*. The remaining eight categories constituted 94.2% of images and were those likely to contain the objects of interest for this paper. Together they should be representative of the variability encountered along the OHL inspection corridors. We sampled from this set of 363,859 images to obtain datasets for our experiments. Each image is 5616×3744 pixels. Figure 1 shows example images illustrative of some of the challenges posed, namely strong background clutter and mixed component types.

B. TOWER CONDITION RATINGS

In this paper we focus on assessing towers based on the condition of their insulators and fittings. Each tower is assessed by inspection of its multiple images. Fittings refers to accessories for attaching insulators to a tower. Specifically, we are interested in insulator U-bolts, which are typically used to attach a suspension insulator to a tower. These components may not have been painted and therefore wear and rust are

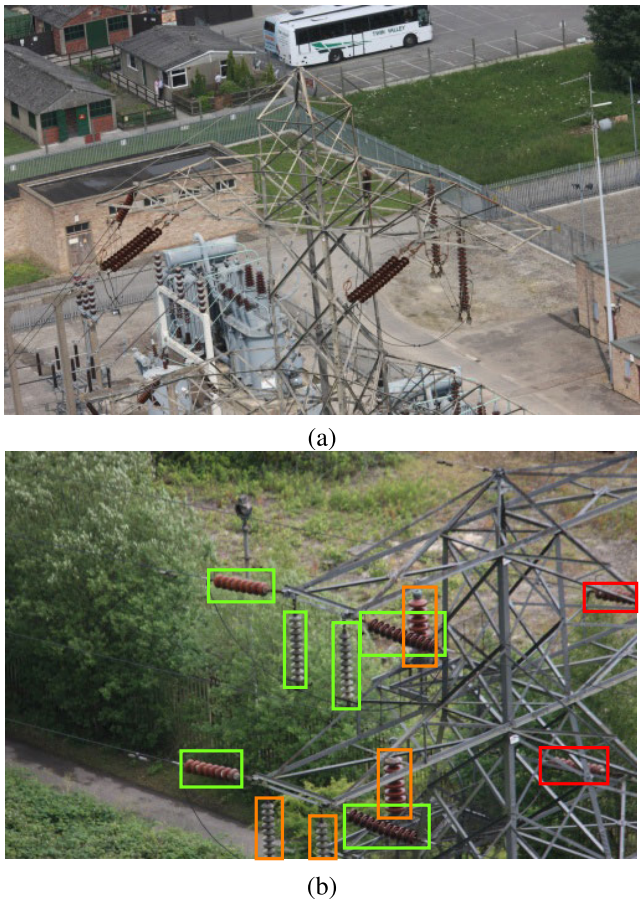


FIGURE 1. **(a) A *top* of tower image with strong background clutter. (b) An *insulator top* image with a mixture of porcelain and glass insulators. Red denotes heavily occluded, orange denotes partial occlusion, and green denotes clear components.**

expected. A critical factor is the severity of wear at the suspension point. Insulator failures can arise from electrical or mechanical faults. Electrical faults are those that reduce or undermine the electrical properties of the device, e.g., cracked or chipped insulator sheds, and in some cases completely missing sheds. Cracks and chips on sheds are typically found on porcelain insulators. Glass insulators would shatter. Mechanical faults arise due to rust resulting from corrosion on the insulator string. Insulator rust can be localised around the cap region or the pin, or in some cases can affect the entire insulator string.

Different assessment criteria were used depending on the type of asset and failure mode being assessed. Table 2 gives the criteria for U-bolts and for mechanical insulator faults. Towers rated *CR-1*, *CR-2*, *CR-3*, *CR-4* or *CR-5* were used in this study whereas those labelled *M*, *N*, *U*, or *'?'* were excluded.

IV. OVERVIEW OF PROPOSED SYSTEM

Figure 2 gives an overview of the process by which a trained system classifies a tower's condition. First, instances of the component type of interest are detected. All detected instances in all images of a tower are cropped and padded to form a bag of instance images. The tower condition is then predicted by a classifier based on this bag. In this paper we focus on insulators and U-bolts as exemplar cases of components of interest. We trained and validated deep instance detectors for these components using disjoint datasets that had been manually annotated with instance bounding boxes.

We should emphasise that tower condition ratings are at the tower level and not at the instance level. It is the tower as a whole that is assigned a condition rating; no condition labels were available in the database for the individual components on which the tower ratings were based. The critical decision threshold lies between *CR-3* and *CR-4*. Towers given ratings of *CR-4* or *CR-5* will be actioned. Therefore, we chose to develop binary classifiers to discriminate between *healthy* towers (*CR-1*, *CR-2*, *CR-3*) and *unhealthy* towers (*CR-4*, *CR-5*).

We compare different methods for developing tower condition classifiers from bags of component instances. A first approach develops an instance classifier by training using class labels transferred from the tower to each of its instances. At test time, a tower is classified by aggregating instance classifications. This approach is likely to work well when the instances on a tower are in similar condition. A second approach employs multiple instance learning (MIL) to train a tower classifier based on the bag of instances. This could be of benefit when the instances on a tower are in varied condition and what matters is the condition of the worst. MIL learning can be resource-intensive so we experiment with a trade-off in which MIL classifiers are trained on sub-bags and their predictions aggregated. Before describing these methods and their evaluation in detail, we first describe the instance detectors.

V. INSTANCE DETECTION

We now detail the methods we used for instance detection before presenting a series of experiments investigating how their performance is affected by object class and instance occlusion.

We tried using Mask R-CNN [46] and RetinaNet [55] for detecting insulators and U-bolts. Mask R-CNN is a two-stage detector based on enhanced Faster R-CNN and incorporates efficient segmentation and localisation in an end-to-end fashion. RetinaNet is a one-stage model that uses focal loss to cope with background-foreground class imbalance. We fine-tuned Mask R-CNN and RetinaNet networks, in both cases with pre-training on the COCO data set [52]. A learning rate of 0.001 was used for the first 10 epochs while fine-tuning only the output layers. The entire network was then trained for between 35 and 50 epochs with a learning rate of 0.0001. Training and testing were carried out in Tensorflow using an RTX2080 machine with 10GB GPU memory. Mini-batch sizes of one and two images were used for Mask R-CNN and RetinaNet, respectively.

We prepared image data for training and testing of instance detectors by annotating bounding boxes for insulators and U-bolts using the VGG Image Annotator (VIA) [59].

TABLE 2. Criteria for rating towers based on the condition of their insulators and U-bolts.

Rating	Insulator mechanical rust	U-bolts or tower attachment
CR-1	OK [Green]	OK [Green]
CR-2	Up to 10% rust. [Green]	Between 10% and 50% of area affected by rust, and/or up to 10% wear [Green]
CR-3	Between 10% and 50% rust. [Orange]	More than 50% of area affected by rust and/or between 10% and 50% wear [Orange]
CR-4	More than 50% rust. [Red]	More than 50% rust AND more than 50% wear indicative of potential failure or already broken [Red]
CR-5	Any single cap appears to be 100% affected by rust, laminated rust, dimpling. Wasting of metal cap, and/or miss-aligned/skewed sheds indicating possible cap/pin seizures. [Red]	Exceptional circumstances [Red]
M	Missing [Red]	Missing [Red]
N	Not applicable	Not applicable
U	Unsure	Unsure
?	Refer to network owner [Red]	Refer to network owner [Red]

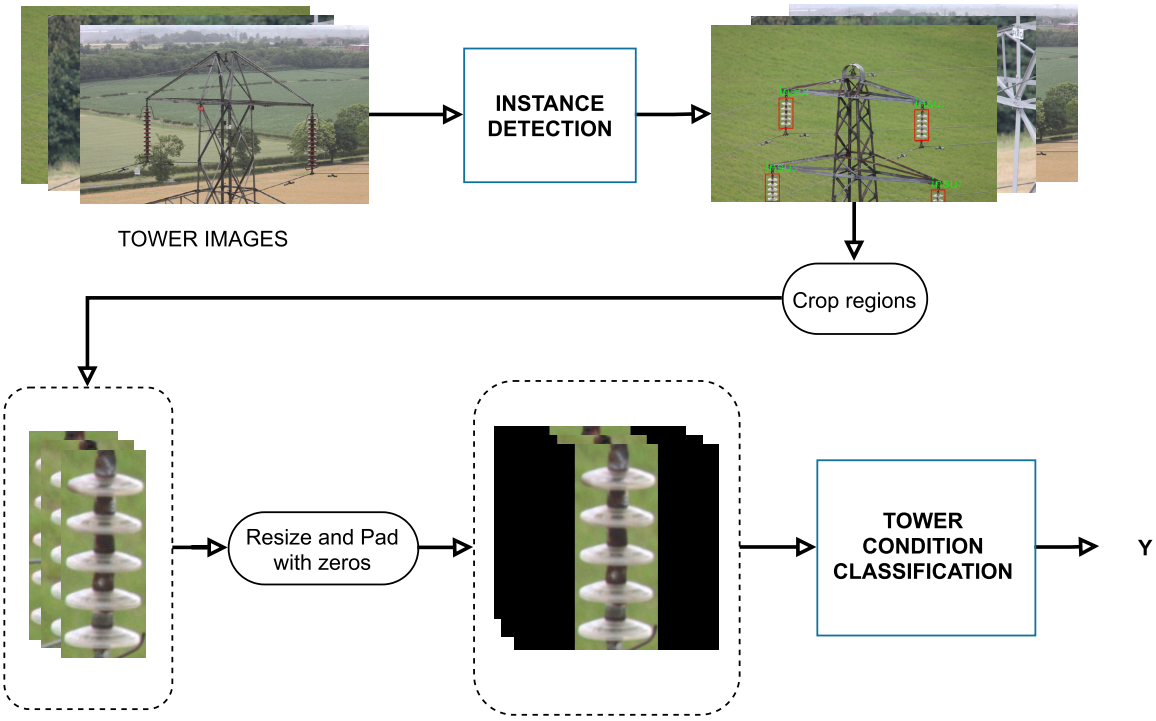


FIGURE 2. Block diagram overview of tower condition classification pipeline.

Each annotated instance was labelled with one of the following occlusion ratings. Figure 1 shows examples of these occlusion levels.

- Clear: 100% within the image with no occlusion.
- Partially occluded: at least half of the object is visible within the image, with some occluded either by the image boundary or by the tower or some other object.
- Heavily occluded: more than half of the object is occluded either by the image boundary or by the tower or some other object.

We evaluated the detection models using average precision (AP) and response time. A ground-truth bounding box is compared with a detected bounding box by computing the ratio of their intersection and their union. Detections with intersection over union (IOU) greater than or equal to 0.5 were considered to be true positives. Precision is computed as $\frac{TP}{TP+FP}$ and recall as $\frac{TP}{TP+FN}$ where TP, FP, and FN denote counts of true positives, false positives, and false negatives, respectively. AP is the area under the precision-recall curve. AP was computed using interpolation as described in [60].

TABLE 3. Effect of occlusion and image size on insulator detection using Mask R-CNN.

Occlusion Levels	Insulator Instances	Average Precision (AP_{50}) 1024x682	512x341
Clear	736	88.7	86.7
Clear + Partial	834	89.4	86.8
Clear + Partial + Heavy	949	83.8	80.8

A. EFFECT OF OCCLUSION AND IMAGE SIZE ON INSULATOR DETECTION

Initially we explored the prevalence of insulators and the complexity of their detection using images sampled at random from the dataset. Specifically, 200 images were sampled from each of the eight viewpoint categories for a total of 1600 images. These were split into 1200 images for training and 400 for testing. In the training set, 367 images contained no insulators while a total of 3,236 insulators were detected in the remaining 833 training images. In the test set there were 949 insulators. Of those, 115 were heavily occluded and a further 98 were partially occluded.

We trained a Mask R-CNN detector using this training set. As shown in Table 3, it obtained an AP of 83.8% overall. This improved to 89.4% if heavily occluded instances were removed. We repeated this experiment with image width and height halved. This reduction resulted in a drop in AP of approximately 3%.

B. DETECTION OF INSULATORS AND U-BOLTS IN HIGH-RISK TOWERS

Inspection reports are usually accompanied by additional comments on major flashpoints; these high-risk cases require immediate attention, such as repairs or replacements, within the inspection window. We investigated detection of components in images of towers deemed to be high risk for the following reasons. Firstly, failure modes are usually underrepresented in inspection data sets; by training detectors on high risk towers, the proportion of faulty examples in our dataset of detected components should increase. Secondly, whereas images taken of high voltage lines from helicopters are often long shots acquired at distance for safety reasons, towers identified as high risk will typically also have close-up shots of suspect components. Such images can serve as an approximation of a UAV-based inspection scenario. Table 4 summarises the high-risk tower dataset consisting of 1,830 images of 748 towers and the training-testing split used. A total of 2,462 insulators and 1,962 U-bolts were annotated in these images.

We compared Mask R-CNN and RetinaNet for insulator and U-bolt detection. Figures 3(a) and (b) show example cropped U-bolts used for training and validation (1024 × 682 pixels). As shown in Table 5, Mask R-CNN and RetinaNet both obtained high AP for insulators. Mask R-CNN obtained a similarly high AP for U-bolts whereas

TABLE 4. Insulators and U-bolts from high-risk towers.

	Towers	Images	Insulator	U-bolt
Training	598	1466	1966	1580
Testing	150	364	496	382
Overall	748	1830	2462	1962

TABLE 5. Detection results in high-risk towers.

Model	Time (ms)	Object	Width	AP_{50}
Mask R-CNN	157	Insulator	1024	96.7
		U-bolt	1024	96.0
Mask R-CNN		Insulator	512	95.6
		U-bolt	512	93.1
RetinaNet	60	Insulator	1024	96.4
		U-bolt	1024	86.9

TABLE 6. RetinaNets for single vs. multi-task detection.

Components	Single (AP_{50})	Multi-task (AP_{50})
Earth-wire U-bolts	90.7	95.1
Insulator U-bolts	96.1	97.9
Insulators	95.3	96.7

RetinaNet did not. Figure 4 shows the precision-recall curves. On inputs of 1024 × 682 pixels, average inference time of 157*ms* and 60*ms* were achieved for Mask R-CNN and RetinaNet, respectively. Reducing input size to 512 × 341 pixels led to a drop in AP (Table 5).

C. FINE-GRAINED AND MULTI-TASK DETECTORS

We investigated a more fine-grained detection setting in which two sub-classes of U-bolt, namely earth-wire U-bolts and insulator U-bolts were distinguished based on the assets they physically support. Furthermore, we compared separately trained detectors to a single multi-task detector. Discriminating between earth-wire and insulator U-bolts requires less tightly cropped images so that sufficient visual context is available. Figures 3(c) and (d) show examples of the crops used. These are contrasted with those in Figures 3(a) and (b) used in the previous experiment. For consistency, we expanded bounding boxes for insulators so that they were also similarly cropped. RetinaNets were trained and tested for single and multi-task detection on the high-risk tower dataset.

As shown in Table 6, multi-task detection achieved better AP results than single-task detectors. Figures 5(a) and 5(b) show precision-recall curves comparing single and multi-task detectors for detection of insulators and insulator U-bolts. The use of expanded bounding boxes appears to have led to improved RetinaNet U-bolt detection.

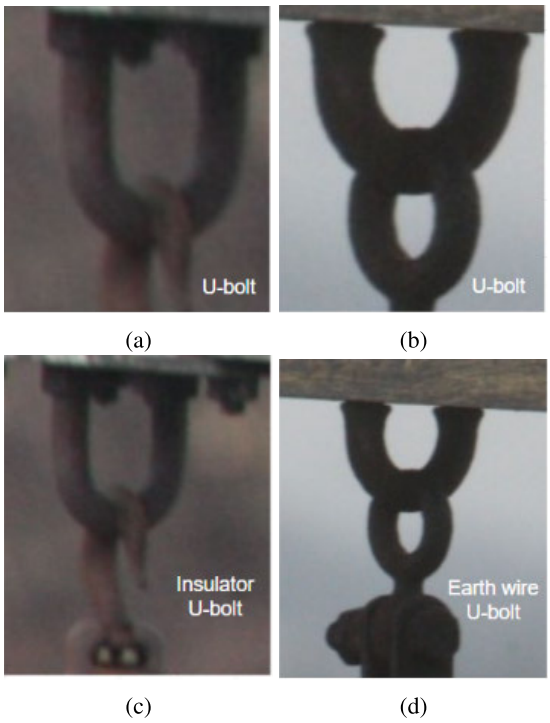


FIGURE 3. Examples of cropped U-bolts used for detection. (a) and (b) are tightly cropped. (c) and (d) are less tightly cropped to facilitate fine-grained discrimination between earth-wire U-bolts and insulator U-bolts.

VI. TOWER CONDITION CLASSIFICATION

The next phase of our inspection pipeline is the classification of towers based on the condition of detected component instances, specifically insulators and U-bolts. As mentioned earlier, two approaches were investigated: (i) aggregating outputs of instance classifiers trained by transferring class labels from towers to instances, and (ii) aggregating outputs of sub-bag classifiers trained using multiple instance learning (MIL).

Instance-based classification used either ResNet50V2 [61] or EfficientNetB0 [62] deep networks with two outputs, trained to classify instances using binary cross-entropy loss. Tower condition labels were used as instance class labels for training.

MIL-based classification used EfficientNetB0 [62] as the base for the multiple instance learning (MIL) network but the top layers were replaced with two fully connected (FC) layers and a final output layer which was a gated attention pooling layer as described in [63]. The pooling is the weighted average of instances where weights are determined by a neural network. Thus the network structure was: feature extraction-based on EfficientNet → FC-64-nodes → Drop-out → FC-32-nodes → Drop-out → Output. This deep MIL network was trainable end-to-end and had approximately 8M parameters.

Due to resource constraints, and so that we could utilise all available instances of a tower, each tower bag was split into sub-bags. We ensured that all sub-bags remained within the same split for either training or testing to avoid any bias. We created sub-bags of up to 8 (MIL_8) and 16 (MIL_{16}) instances. Sub-bag class predictions were aggregated at test time.

TABLE 7. Dataset breakdown for experiments on tower classification in terms of the condition of their insulators.

Class	Towers	Sample Towers	Sample Images	Detected Insulators
Healthy	4,048	600	15,216	49,290
Unhealthy	1,253	400	10,588	32,933
Total	5,301	1,000	25,804	82,223

Networks were trained from scratch using mini-batch stochastic gradient descent with a mini-batch size of 8 instances. The learning rate was initially $\alpha_0 = 0.001$, and was scaled based on Eqn. (1) where τ is the current epoch and τ_{max} is the maximum number of epochs.

$$\alpha = \alpha_0(1 - \frac{\tau}{\tau_{max}}) \tag{1}$$

Data augmentation consisted of flipped training images horizontally at random. Models were evaluated using 10-fold cross-validation. We compare the use of mean and max operators to determine the tower class label $\hat{Y}$ from the predicted sub-bag or instance labels.

VII. RESULTS

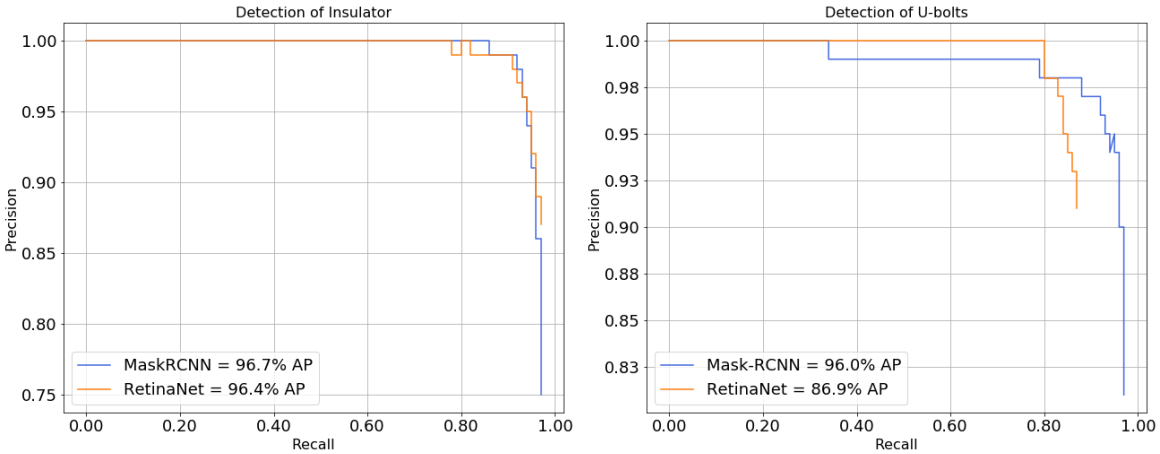
A. TOWER CLASSIFICATION BASED ON INSULATORS

We experimented with classification of towers as healthy (*CR-1*, *CR-2*, *CR-3*) or unhealthy (*CR-4*, *CR-5*) in terms of the condition of their insulators. We created a dataset by sampling 1,000 towers, 200 per condition rating (CR), from the 5,301 towers that were not used in the insulator detection experiments (Section V). These 1,000 towers had a total of over 25k images in which more than 82k instances (sub-images) of insulators were automatically detected. Table 7 gives a more detailed breakdown of these data. Figure 6 shows examples of detected insulator instances. All detected instances were used in what follows, regardless of whether they were of glass, porcelain, composite or ceramic insulators, whether they were from suspension or tension towers, or indeed whether they were false positive detections. Each detected instance sub-image was scaled and cropped to 224 × 224 pixels.

Figure 7 shows receiver operating characteristic (ROC) curves for tower classification, Table 8 summarises results in terms of area under the ROC curve (AUC), and Table 9 gives confusion matrices at operating points selected to maximise Youden's index [64].

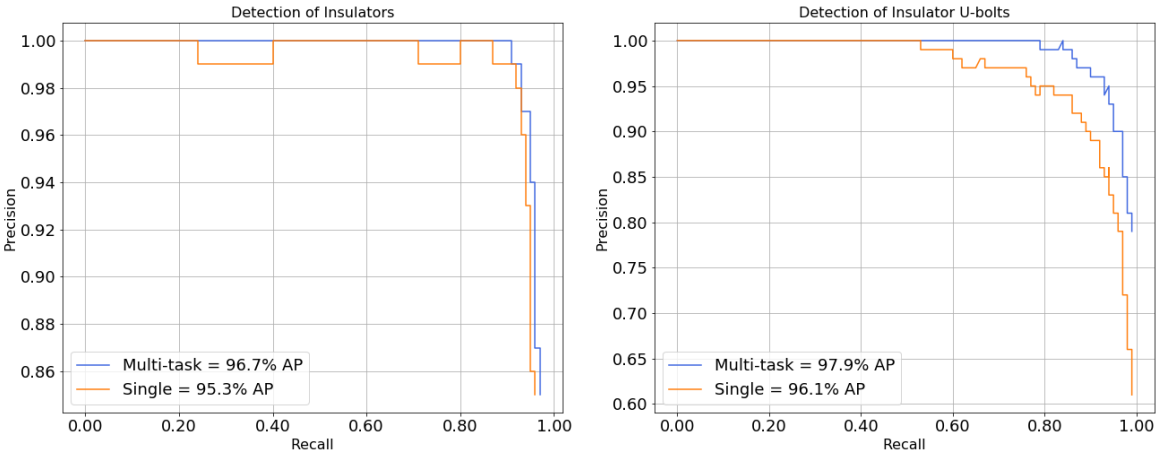
1) TOWER CLASSIFICATION BY AGGREGATING INSTANCE CLASSIFICATIONS

The ROC curve obtained by averaging EfficientNetB0 instance classification scores dominated that obtained by taking the maximum instance classification score (Figure 7(a)); averaging gave an AUC of 0.94 whereas taking the maximum



(a) Insulator detection (b) U-bolt detection

FIGURE 4. Mask R-CNN and RetinaNet on high-risk towers.



(a) Insulator detection (b) Insulator U-bolt detection

FIGURE 5. RetinaNet for single and multi-task detections.

gave a lower AUC of 0.90. This compared favourably with the larger ResNet50V2 network which had lower AUCs of 0.90 and 0.81, respectively. As shown in Table 9, 1,000 towers were classified and EfficientNet performed better than other classifiers with 864 (healthy and unhealthy) towers correctly classified.

2) TOWER CLASSIFICATION BY AGGREGATING MIL SUB-BAG CLASSIFICATIONS

Figure 7(b) shows ROC curves obtained by aggregating MIL sub-bag classifications using sub-bags of 16 instances. There was no clear distinction between aggregation using mean and max. Table 8, we show network parameters and AUC values of MIL-based classifiers. Increasing the number of instances in sub-bags from 8 to 16 improved AUC from 0.90 to 0.92 (in the case of mean aggregation) but did not do better than aggregating instance classifications. As shown in shown in Table 9, MIL_{16} followed by averaging sub-bags predictions performed better than that trained with sub-bags of 8 instances.

TABLE 8. Summary of tower classification results based on insulators and U bolts.

Method	Network Param.	Insulator-based Mean	Insulator-based Max	Insulator U-bolt Mean	Insulator U-bolt Max
EfficientNetB0	4M	0.94	0.90	0.91	0.81
ResNet50V2	23M	0.90	0.81	-	-
MIL_{16}	8M	0.92	0.91	0.89	0.89
MIL_8	8M	0.90	0.88	-	-

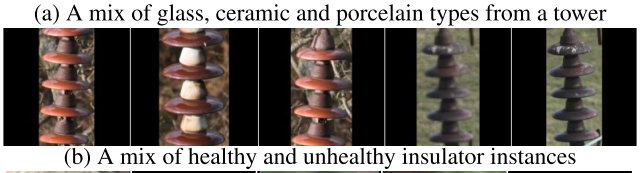
B. TOWER CLASSIFICATION BASED ON U-BOLTS

We experimented with classification of towers as healthy (*CR-1*, *CR-2*, *CR-3*) or unhealthy (*CR-4*, *CR-5*) in terms of the condition of their insulator U-bolts. We created a dataset for this purpose using a sample of the suspension (S-type) towers.

TABLE 9. Insulator-based tower classification.

		ResNet50V2		EfficientNetB0		MIL_8		MIL_{16}	
		Healthy	Unhealthy	Healthy	Unhealthy	Healthy	Unhealthy	Healthy	Unhealthy
Mean	Healthy	516	84	506	94	493	107	488	112
	Unhealthy	75	325	42	358	70	330	50	350
Max	Healthy	460	140	487	113	520	80	462	138
	Unhealthy	110	290	46	354	109	291	41	359

(a) A mix of glass, ceramic and porcelain types from a tower



(b) A mix of healthy and unhealthy insulator instances

(c) False positive insulator detections

FIGURE 6. Examples of detected insulator instances used for tower classification.

Only 323 of the 3,647 S-type towers are categorised as unhealthy (Table 10) so to create balanced dataset we sampled 320 healthy and 320 unhealthy towers. Four towers (1 healthy and 3 unhealthy) were excluded by the detector as not having insulator U-bolts. The remaining 636 towers had a total of over 15k images in which more than 19k instances (sub-images) of insulator U-bolts were automatically detected. Table 11 gives a more detailed breakdown of these data. Examples of detected insulator U-bolts are shown in Figure 10(a). On average, 30 U-bolt instances were detected per tower; some had as many as 90. These included some false positives as in Figure 10(b). All detected instances were used in what follows, including false positive detections. Each detected instance sub-image was scaled, cropped and zero-padded to 224 × 224 pixels.

TABLE 10. Numbers of suspension towers, their images, and insulator U-bolt instances detected.

Class	Towers	Images	Detected Towers	Detected U-bolts
Healthy	3,324	87,501	3,315	97,128
Unhealthy	323	7,118	320	10,151
Total	3,647	94,619	3,635	107,279

On the entire 3,647 towers, the detector excluded 12 towers as not having insulator U-bolts. Nine of these towers have pin insulator types installed, which are not supported by U-bolts. An example of a tower with pin-type insulators is shown in Figure 8(b); our detector was not trained to detect pins.

TABLE 11. Dataset used for tower classification by U-bolt condition.

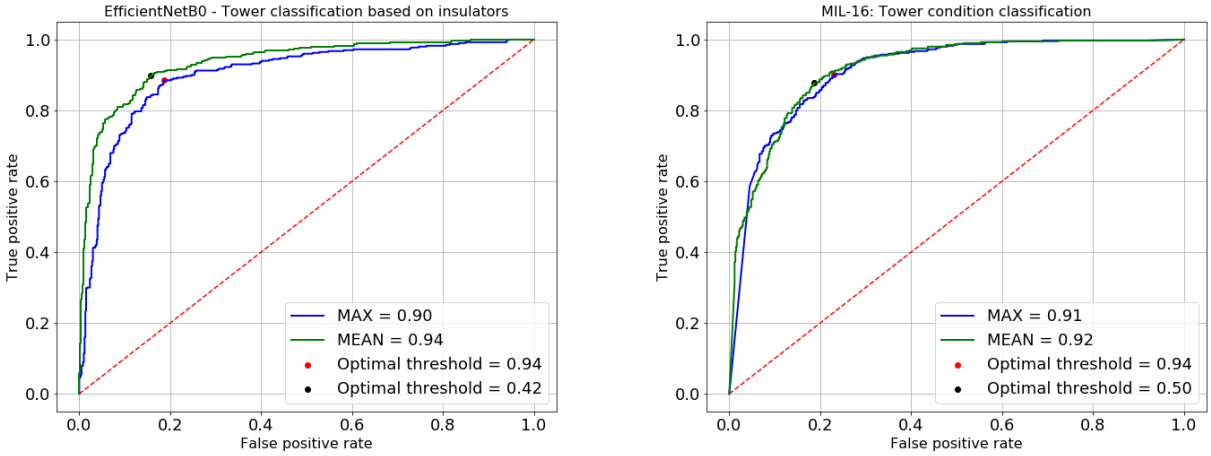
Class	Towers	Images	Insulator U-bolts
Healthy	319	8,653	9,168
Unhealthy	317	7,024	10,083
Total	636	15,677	19,251

Furthermore, the T-type tower shown in Figure 8(a) had been wrongly labelled as S-type; we detected U-bolts from S-types only so this erroneously labelled image was not applicable. Figure 9(a) and Figure 9(b) are of interest as these are S-type towers. All insulator U-bolts in Figure 9(a) were heavily occluded and were missed by the detector. All the images of the tower in Figure 9(b) were taken at long range. As a result, the size of the components was always too small to be identified by the detector. These examples illustrate the limitations of the detector in terms scale and level of occlusion.

Table 8 gives AUC results. Aggregating MIL sub-bag classifications using sub-bags of 16 instances achieved an AUC of 0.89 irrespective of whether mean or max was used for aggregation. Aggregating instance classification did best when using mean aggregation, achieving an AUC of 0.91, but poorly when using max aggregation, dropping to an AUC of 0.81.

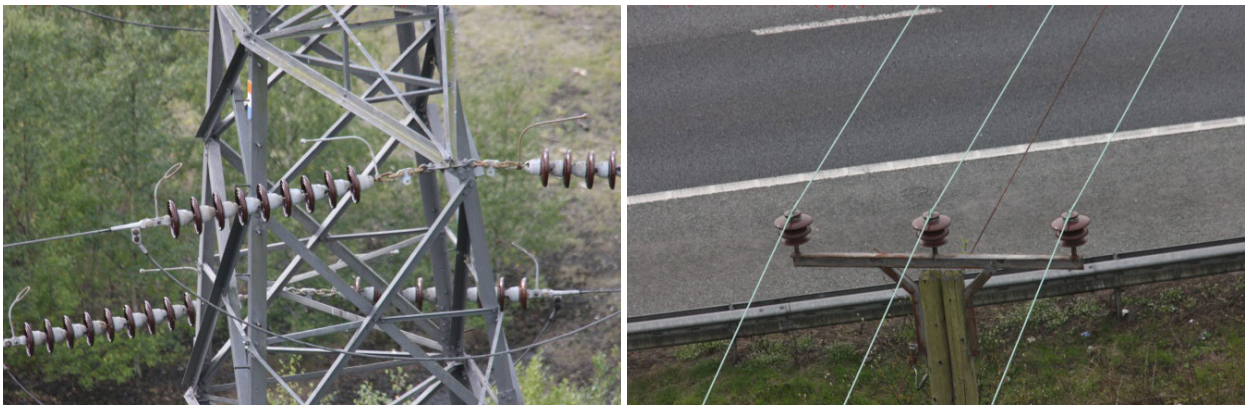
Table 12 shows confusion matrices obtained at operating points chosen to maximise Youden's index. In particular, MIL with max aggregation over sub-bags resulted in 541 (266 healthy and 275 unhealthy) towers being correctly classified and 95 (53 healthy and 42 unhealthy) towers being mis-classified. Mean aggregation of instance classifications resulted in 534 (260 healthy and 274 unhealthy) towers being correctly classified and 102 (59 healthy and 43 unhealthy) towers being mis-classified. Table 13 expands the confusion matrices to show how original ratings (CR1, CR2, CR3, CR4, & CR5) were classified as healthy and unhealthy. As would be expected, most confusion occurs between CR-3 and CR-4, whereas CR-1 and CR-5 were classified with few errors.

Given the promising results on the balanced data sample, we further evaluated aggregation of instance classification using all 3,647 towers (of which 3,315 healthy and 320 unhealthy towers resulted in insulator U-bolt detections). As shown in Table 10, 97,128 U-bolt instances were detected in healthy S-type towers and 10,151 in unhealthy S-type towers. This is a highly unbalanced dataset so class weighting of 1:10 was applied to the loss function during training.



(a) Insulator-based tower classification (b) MIL_{16}-based tower classification

FIGURE 7. ROC curves for Instance-based and MIL-based tower condition classification.



(a) T-type tower wrongly labelled as S-type (Human error) (b) Tower with pin insulators not applicable for insulator U-bolts

FIGURE 8. Examples of data set limitations.

(a) Tower with occluded insulator U-bolts (b) Tower in long-range image

FIGURE 9. Examples of system limitations.

Figure 11 shows ROC curves obtained using mean and max aggregation; the curve for mean aggregation dominates. The classifiers achieved AUCs of 0.98 and 0.94 for mean and max aggregation, respectively. Table 14 gives confusion matrices, and Table 15 the expanded confusion matrices showing how the five tower condition ratings get classified as healthy

TABLE 12. Confusion matrices for tower classification based on U-bolts.

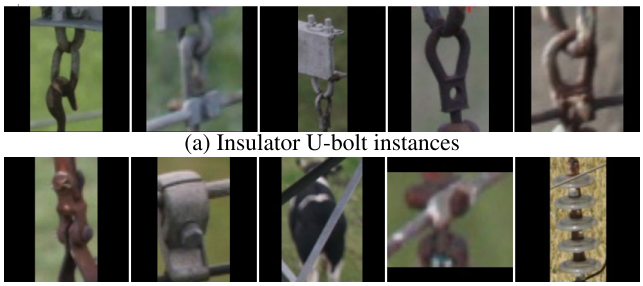
		Instance-based		MIL_{16}	
		Healthy	Unhealthy	Healthy	Unhealthy
Mean	Healthy	260	59	269	50
	Unhealthy	43	274	49	268
Max	Healthy	263	56	266	53
	Unhealthy	104	213	42	275

TABLE 13. Expanded confusion matrices for tower classification based on U-bolts (using mean aggregation).

		Instance-based		MIL_{16}	
Class	CR	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	80	2	81	1
Healthy	2	76	4	76	4
Healthy	3	104	53	112	45
Unhealthy	4	38	208	44	202
Unhealthy	5	5	66	5	66

TABLE 14. Confusion matrices for tower classification based on U-bolts (using aggregation of instance classifications on the unbalanced dataset).

	Instance-based (Mean)		Instance-based (Max)	
	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	3191	124	2937	378
Unhealthy	31	289	37	283



(a) Insulator U-bolt instances

(b) False positives of insulator U-bolt detection

FIGURE 10. Sub-images extracted as insulator U-bolts.

and unhealthy. Using the averaging instance classifications method, 3,191 out of 3,315 healthy towers and 289 out of 320 unhealthy towers were correctly classified. A total of 155 (31 unhealthy and 124 healthy) towers were missed. Four of the 72 CR-5 towers (6%) were mis-classified as healthy and just one of the 865 CR-1 towers (0.1%) was mis-classified as unhealthy.

VIII. DISCUSSION

This study leveraged real-world data from several years of inspection with a wide variety of component types, failure modes and views. Aerial images of OHL towers have highly cluttered backgrounds which, together with variability in terms of component types and scale, makes their analysis challenging. The data were representative of this complexity, covering a diversity of clutter (e.g., roof tops, vegetation) and

TABLE 15. Expanded confusion matrices for tower classification based on U-bolts (using aggregation of instance classifications on the unbalanced dataset).

		Instance-based (Mean)		Instance-based (Max)	
Class	CR	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	864	1	841	24
Healthy	2	840	7	797	50
Healthy	3	1487	116	1299	304
Unhealthy	4	27	221	34	214
Unhealthy	5	4	68	3	69

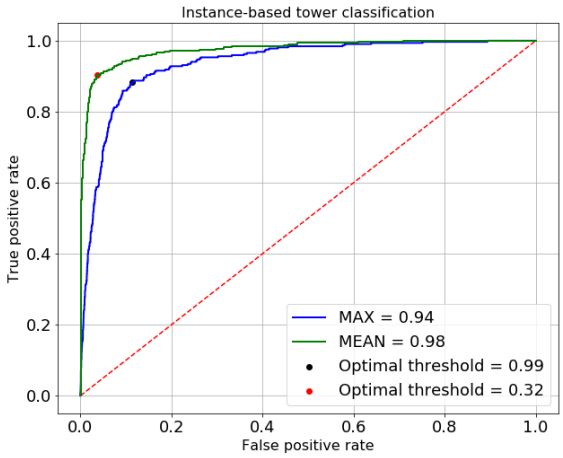


FIGURE 11. ROC curves for tower classification based on U-bolt condition (using aggregation of instance classifications on the unbalanced dataset).

component types, which should be beneficial for generalisation. As noted in [6], [50], insulators of diverse shapes, lengths and orientations may be present on a single transmission tower. Whereas some previous studies have focused on detecting glass insulators [31] or porcelain and composite insulators [30], [49], this study detected four insulator types simultaneously, i.e., porcelain, ceramic, glass and composite. Furthermore, we applied deep learning object detection methods to the detection of U-bolts, and we reported experiments investigating the effects of occlusion, scale, and multi-task detection.

Components that were missed tended to be those that were distant, or that were heavily occluded by tower structures and image boundaries. However, it is likely that other viewpoints would result in their detection. Discarding heavily occluded insulator instances improved Mask R-CNN performance by over 6% from 83.8% AP to 89.4% AP (Table 3). Scaling input resolution down by half degraded performance by about 3%. This performance is consistent with the assertion that close-up images are necessary for electricity tower inspection [1]. When tested on close-up images from high-risk towers, Mask R-CNN achieved an impressive 96.7% AP. Mask R-CNN was more robust for detecting small U-bolts. For the larger insulators, RetinaNet had comparable AP with a run-time 3 times faster. Whereas training using tightly fitting bounding boxes worked well for insulators, expanded bounding boxes

were more effective for detecting the two types of U-bolt. Multi-task RetinaNet detection performed better than detecting objects separately.

We classified towers as healthy or unhealthy, in terms of their insulators or U-bolts, based on all detected component instances. We explored deep learning methods for multiple instance classification and two methods for aggregating bag predictions. Our results show similar performances using mean and max as aggregation operators with the mean operator achieving slightly better results on both insulators and U-bolts.

OHL network assets are inspected periodically, and in an operational setting it is likely to be more costly for an automated system to misclassify unhealthy towers than to misclassify healthy towers during an inspection window. Classification thresholds can be set accordingly to reflect these asymmetric costs. For example, the ROC curve in Figure 11 suggests that 97.5% of towers with unhealthy U-bolts would be flagged for human inspection with a corresponding 20% false positive rate.

IX. CONCLUSION AND FUTURE WORK

Electricity networks overhead Line (OHL) assets represent critical infrastructure which must be supported by cost effective asset management. New demands as a result of the increasing end-user energy demand e.g., decarbonisation of transport and the increasingly complex energy flows from distributed generation are driving the needs for advancements in current technology and practices. Automating the inspection of OHL assets is a challenging task requiring multiple image processing steps to detect the components of interest and their failure modes from cluttered backgrounds. Deep learning can be effective at learning the complex representations needed for such tasks, but a challenge with deep learning is the need for sufficient numbers of properly labelled images to achieve good generalisation. Such data requires significant additional investment to create. We leveraged a robust inspection dataset which is representative of the diverse asset types and failure modes in real-world scenarios. Assets were automatically extracted from multiple images of a tower. The towers were then classified as healthy or unhealthy based on the detected assets. We introduced the automated detection and condition analysis of U-bolts. Our tower classification results are consistent with the original tower condition ratings, demonstrating the effectiveness of deep learning for prognostic activities along overhead line networks from aerial images. We have demonstrated that tower condition classifiers can be trained effectively without labelling the condition ratings of individual components (insulators and U-bolts) upon which such classifications are ultimately based.

Current inspection pipelines for electrical line surveys capture aerial images from helicopters. However, the literature reviewed shows a trend towards the replacement of helicopters with safer, more flexible and cost effective aerial platforms such as unmanned aerial vehicles (UAVs). The methods presented in this paper were based on tower images taken from helicopters. There could be some performance issues when deployed on a different platform because of the variation in image quality. Future work should focus on the verification of the proposed pipeline for unmanned aerial vehicle surveys. This pipeline could also be used for the assessment of tower paintwork degradation. There are other failure modes that need to be identified such as trees in breach of power-lines. Tree cutting accounts for a sizeable proportion of investment in OHL network maintenance.

ACKNOWLEDGMENT

Northern Powergrid provided data for this study.

REFERENCES

[1] D. I. Jones and G. K. Earp, "Camera sightline pointing requirements for aerial inspection of overhead power lines," *Electr. Power Syst. Res.*, vol. 57, no. 2, pp. 73–82, Mar. 2001.

[2] L. Matikainen, M. Lehtomäki, E. Ahokas, J. Hyyppä, M. Karjalainen, A. Jaakkola, A. Kukko, and T. Heinonen, "Remote sensing methods for power line corridor surveys," *ISPRS J. Photogramm. Remote Sens.*, vol. 119, pp. 10–31, Sep. 2016.

[3] R. Eyre-Walker, G. Howarth, R. Ahmed, and J. Lewin, "Application of advanced condition assessment and asset management techniques on steel tower overhead line electricity networks," in *Proc. Asset Manage. Conf.*, 2014, pp. 1–6.

[4] R. Fuentes, T. Chapman, M. Cook, J. Scanlan, Z. Li, and R. C. Richardson, "Briefing: UK-RAS white paper in robotics and autonomous systems for resilient infrastructure," *Proc. Inst. Civil Eng. Smart Infrastruct. Construct.*, vol. 170, no. 3, pp. 72–79, Sep. 2017.

[5] X. Li, H. Su, and G. Liu, "Insulator defect recognition based on global detection and local segmentation," *IEEE Access*, vol. 8, pp. 59934–59946, 2020.

[6] C. Sampedro, J. Rodriguez-Vazquez, A. Rodriguez-Ramos, A. Carrio, and P. Campoy, "Deep learning-based system for automatic recognition and diagnosis of electrical insulator strings," *IEEE Access*, vol. 7, pp. 101283–101308, 2019.

[7] H. H. Alhelou, M. Hamedani-Golshan, T. Njenda, and P. Siano, "A survey on power system blackout and cascading events: Research motivations and challenges," *Energies*, vol. 12, no. 4, p. 682, Feb. 2019.

[8] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1922–1928, Nov. 2005.

[9] M. Schmidthaler and J. Reichl, "Assessing the socio-economic effects of power outages ad hoc: An application of BLACKOUT-SIMULATOR.Com covering 266 European regions, 9 economic sectors and households separately," *Comput. Sci. Res. Develop.*, vol. 31, no. 3, pp. 157–161, Aug. 2016.

[10] R. Gonzales, "California officials blame PG&E for state's deadliest wildfire: NPR," Nat. Public Radio, Washington, DC, USA, Tech. Rep., May 2019. [Online]. Available: https://www.npr.org/2019/05/15/723753237/pg-e-transmission-lines-caused-californias-deadliest-wildfire-state-officials-sa?t=1615304407590

[11] A. Thompson, B. Kazemtabrizi, C. J. Crabtree, C. Dao, F. Dinmohamadi, and D. Flynn, "Reliability and economic evaluation of high voltage direct current interconnectors for large-scale renewable energy integration and transmission," in *Proc. 15th IET Int. Conf. AC DC Power Transmiss. (ACDC)*, 2019, pp. 1–6.

[12] D. Roman, R. Dickie, D. Flynn, and V. Robu, "A review of the role of prognostics in predicting the remaining useful life of assets," in *Proc. 27th Eur. Saf. Rel. Conf.*, Jun. 2017, p. 135.

[13] E. Miguelañez-Martin and D. Flynn, "Embedded intelligence supporting predictive asset management in the energy sector," in *Proc. Asset Manage. Conf.*, 2015, pp. 1–7.

[14] A. Odo, S. McKenna, D. Flynn, and J. Vorstius, "Towards the automatic visual monitoring of electricity pylons from aerial images," in *Proc. 15th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, Feb. 2020, pp. 566–573.

[15] J. Hao, H. Wulin, C. Jing, L. Xinyu, M. Xiren, and Z. Shengbin, "Detection of bird nests on power line patrol using single shot detector," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2019, pp. 3409–3414.
[16] Z. Qiu, X. Zhu, D. Shi, and Y. Kuang, "Recognition of transmission line related bird species based on image feature extraction and support vector machine," in *Proc. IEEE Int. Conf. High Voltage Eng. Appl. (ICHVE)*, Sep. 2020, pp. 1–4.
[17] Q. Guo and X. Hu, "Power line icing monitoring method using binocular stereo vision," in *Proc. 12th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2017, pp. 1905–1908.
[18] Q. Guo, J. Xiao, and X. Hu, "New keypoint matching method using local convolutional features for power transmission line icing monitoring," *Sensors*, vol. 18, no. 3, p. 698, Feb. 2018.
[19] J. Li, Q. Shao, K. Xue, C. Wang, and W. Hu, "The icing-thickness detection of high-voltage transmission line based on machine vision," in *Proc. IEEE Int. Conf. Inf. Autom. (ICIA)*, Jul. 2017, pp. 381–385.
[20] X.-B. Huang, J.-Q. Li, Y. Zhang, and F. Zhang, "Recognition and detection technology of ice-covered insulators under complex environment," *High Voltage Eng.*, vol. 43, no. 3, pp. 891–899, Mar. 2017.
[21] X. B. Huang, F. Zhang, H. Li, and X. Liu, "An online technology for measuring icing shape on conductor based on vision and force sensors," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3180–3189, Dec. 2017.
[22] S. Rong and L. He, "A joint faster RCNN and stereovision algorithm for vegetation encroachment detection in power line corridors," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2020, pp. 1–5.
[23] Y. Zhang, X. Huang, J. Jia, and X. Liu, "A recognition technology of transmission lines conductor break and surface damage based on aerial image," *IEEE Access*, vol. 7, pp. 59022–59036, 2019.
[24] Y. Pan, F. Liu, J. Yang, W. Zhang, Y. Li, C. S. Lai, X. Wu, L. L. Lai, and B. Hong, "Broken power strand detection with aerial images: A machine learning based approach," in *Proc. IEEE Int. Smart Cities Conf. (ISC2)*, Sep. 2020, pp. 1–7.
[25] S. Fang, L. Sheng, and W. Xiaoyu, "Detection method of transmission line broken stock defects in aircraft inspection based on image processing technology," in *Proc. IEEE 4th Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Jun. 2020, pp. 1715–1719.
[26] J. Oh and C. Lee, "3D power line extraction from multiple aerial images," *Sensors*, vol. 17, no. 10, p. 2244, Sep. 2017.
[27] H. Wang, S. Han, L.-J. Lv, and L.-J. Jin, "Transmission line sag measurement based on single aerial image," in *Proc. 24th Int. Conf. Mechatronics Mach. Vis. Pract. (M2VIP)*, Nov. 2017, pp. 1–5.
[28] X. Liu, X. Miao, H. Jiang, and J. Chen, "Review of data analysis in vision inspection of power lines with an in-depth discussion of deep learning technology," 2020, *arXiv:2003.09802*.
[29] X. Huang, H. Zhang, and Y. Zhang, "Automatic identification and location technology of glass insulator self-shattering," *J. Electron. Imag.*, vol. 26, no. 6, Nov. 2017, Art. no. 063014.
[30] F. Gao, J. Wang, Z. Kong, J. Wu, N. Feng, S. Wang, P. Hu, Z. Li, H. Huang, and J. Li, "Recognition of insulator explosion based on deep learning," in *Proc. 14th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2017, pp. 79–82.
[31] Z. Ling, D. Zhang, R. C. Qiu, Z. Jin, Y. Zhang, X. He, and H. Liu, "An accurate and real-time method of self-blast glass insulator location based on faster R-CNN and U-Net with aerial images," *CSEE J. Power Energy Syst.*, vol. 5, no. 4, pp. 474–482, Dec. 2019.
[32] Y. Yang, L. Wang, Y. Wang, and X. Mei, "Insulator self-shattering detection: A deep convolutional neural network approach," *Multimedia Tools Appl.*, vol. 78, no. 8, pp. 10097–10112, Apr. 2019.
[33] V. S. Murthy, K. Tarakanath, D. K. Mohanta, and S. Gupta, "Insulator condition analysis for overhead distribution lines using combined wavelet support vector machine (SVM)," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 17, no. 1, pp. 89–99, Feb. 2010.
[34] Z. A. Siddiqui, U. Park, S.-W. Lee, N.-J. Jung, M. Choi, C. Lim, and J.-H. Seo, "Robust powerline equipment inspection system based on a convolutional neural network," *Sensors*, vol. 18, no. 11, p. 3837, Nov. 2018.
[35] S. Fang, Z. Mingze, L. Sheng, W. Xiaoyu, and C. Haiyang, "Fast detection method of insulator fault based on image processing technology," in *Proc. IEEE 5th Inf. Technol. Mechatronics Eng. Conf. (ITOEC)*, Jun. 2020, pp. 400–406.
[36] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 4, pp. 1486–1498, Apr. 2020.
[37] L. Xin, H. Jin, Y. Tu, Z. Yuan, Z. Lv, and C. Wang, "Defect detection and characterization of RTV silicone rubber coating on insulator based on visible spectrum image," *IEEE Trans. Power Del.*, vol. 35, no. 6, pp. 2734–2736, Dec. 2020.
[38] B. Jalil, G. R. Leone, M. Martinelli, D. Moroni, M. A. Pascali, and A. Berton, "Fault detection in power equipment via an unmanned aerial system using multi modal data," *Sensors*, vol. 19, no. 13, p. 3014, Jul. 2019.
[39] M. Oberweger, A. Wendel, and H. Bischof, "Visual recognition and fault detection for power line insulators," in *Proc. 19th Comput. Vis. Winter Workshop*, 2014, pp. 1–8.
[40] Y. Zhai, D. Wang, M. Zhang, J. Wang, and F. Guo, "Fault detection of insulator based on saliency and adaptive morphology," *Multimedia Tools Appl.*, vol. 76, no. 9, pp. 12051–12064, May 2017.
[41] S. P. Potnuru and P. R. Bhima, "Image processing and machine learning applied for condition monitoring of 11-kV power distribution line insulators using curvelet and LTP features," in *Proc. IEEE Int. Conf. Power, Control, Signals Instrum. Eng. (ICPCSI)*, Sep. 2017, pp. 3012–3017.
[42] K. Zhang and L. Yang, "Insulator segmentation algorithm based on k-means," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2019, pp. 4747–4751.
[43] Z. Hong-Bin, H. Long, and L. Yun-Feng, "Target tracking method of transmission line insulator based on multi feature fusion and adaptive scale filter," in *Proc. 5th Asia Conf. Power Electr. Eng. (ACPEE)*, Jun. 2020, pp. 1626–1630.
[44] Y. Yu, H. Cao, Z. Wang, Y. Li, K. Li, and S. Xie, "Texture-and-shape based active contour model for insulator segmentation," *IEEE Access*, vol. 7, pp. 78706–78714, 2019.
[45] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
[46] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE ICCV*, Dec. 2017, pp. 2980–2988.
[47] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Computer Vision—ECCV*. Cham, Switzerland: Springer, 2016, pp. 21–37.
[48] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 6517–6525.
[49] X. Miao, X. Liu, J. Chen, S. Zhuang, J. Fan, and H. Jiang, "Insulator detection in aerial images for transmission line inspection using single shot multibox detector," *IEEE Access*, vol. 7, pp. 9945–9956, 2019.
[50] H. Ohta, Y. Sato, T. Mori, K. Takaya, and V. Kroumov, "Image acquisition of power line transmission towers using UAV and deep learning technique for insulators localization and recognition," in *Proc. 23rd Int. Conf. Syst. Theory, Control Comput. (ICSTCC)*, Oct. 2019, pp. 125–130.
[51] V. N. Nguyen, R. Jenssen, and D. Roverso, "Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning," *Int. J. Elect. Power Energy Syst.*, vol. 99, pp. 107–120, Jul. 2018.
[52] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Computer Vision—ECCV*. Cham, Switzerland: Springer, 2014, pp. 740–755.
[53] H. Chen, Z. He, B. Shi, and T. Zhong, "Research on recognition method of electrical components based on YOLO V3," *IEEE Access*, vol. 7, pp. 157818–157829, 2019.
[54] J. Han, Z. Yang, Q. Zhang, C. Chen, H. Li, S. Lai, G. Hu, C. Xu, H. Xu, D. Wang, and R. Chen, "A method of insulator faults detection in aerial images for high-voltage transmission lines inspection," *Appl. Sci.*, vol. 9, no. 10, p. 2009, May 2019.
[55] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 318–327, Feb. 2020.
[56] J. Wang, K. Wang, G. Liu, W. Zhou, and Z. Zhou, "Recognition of defects in pins based on generative adversarial network and RetinaNet," *Huanan Ligong Daxue Xuebao/J. South China Univ. Technol. Natural Sci.*, vol. 48, no. 2, pp. 1–8, 2020.
[57] Z. Zhou, G. Yuan, W. Feng, S. Gu, and P. Fan, "Target recognition and evaluation of typical transmission line equipment based on deep learning," in *Proc. PURPLE MOUNTAIN FORUM Int. Forum Smart Grid Protection and Control* (Lecture Notes in Electrical Engineering), vol. 585. Singapore: Springer, 2020, pp. 701–709.

[58] A. Jiang, N. Yan, B. Shen, C. Gu, H. Zhu, and H. Huang, "Research on infrared image recognition method of power equipment based on deep learning," in *Proc. IEEE Int. Conf. High Voltage Eng. Appl. (ICHVE)*, Sep. 2020, pp. 1–4.
[59] A. Dutta and A. Zisserman, "The VIA annotation software for images, audio and video," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 2276–2279.
[60] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," in *Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Jul. 2020, pp. 237–242.
[61] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Computer Vision—ECCV*. Cham, Switzerland: Springer, 2016, pp. 630–645.
[62] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Mach. Learn. Res.*, vol. 97, K. Chaudhuri and R. Salakhutdinov, Eds. Long Beach, CA, USA: PMLR, Jun. 2019, pp. 6105–6114.
[63] M. Ilse, J. Tomczak, and M. Welling, "Attention-based deep multiple instance learning," in *Proc. 35th Int. Conf. Mach. Learn.*, vol. 80. Stockholm, Sweden: Stockholmsmässan, Jul. 2018, pp. 2127–2136.
[64] N. Smits, "A note on Youden's J and its cost ratio," *BMC Med. Res. Methodol.*, vol. 10, no. 1, p. 89, Dec. 2010.

ANICETUS ODO received the B.Eng. degree in computer science and engineering from the Enugu State University of Science and Technology, Enugu State, Nigeria, in 2001, and the M.Eng. degree in electronics and computer engineering from Nnamdi Azikiwe University, Anambra, Nigeria, in 2009. He is currently pursuing the Ph.D. degree with the University of Dundee, U.K. From 2007 to 2012, he started his career as a Technologist and later moved on to lecturing at the Computer Engineering Department, Enugu State University of Science and Technology, from 2012 to 2017. He is a member of the Computer Vision and Image Processing (CVIP) Group. In 2017, he received a Research Grant from the Tertiary Education Trust Fund (TETFUND), Nigeria.

STEPHEN MCKENNA received the B.Sc. degree (Hons.) in computer science from The University of Edinburgh, in 1990, and the Ph.D. degree in medical image analysis from the University of Dundee, in 1994. He was an EU Research Fellow, Italy, from 1994 to 1995, and an EPSRC Postdoctoral Researcher at the Queen Mary University of London, from 1995 to 1998. He has held visiting research positions at George Mason University, BT, USA, and Universidad Iberoamericana. He is currently a Professor (a Personal Chair) and the Computing Research Lead at the University of Dundee. He has involved interdisciplinary collaborations with biologists, surgeons, pathologists, oncologists, veterinarians, psychologists, artists, designers, charities, museums, and small and large companies. He teaches courses on machine learning, image analysis, and their application to healthcare. His research interests include biomedical image analysis, computer vision, and machine learning. He serves on journal editorial boards which include *Machine Vision and Applications*, *Journal of Imaging*, and *Scientific Data*.

DAVID FLYNN (Member, IEEE) received the B.Eng. degree (Hons.) in electrical and electronic engineering, the M.Sc. degree (Hons.) in microsystems, and the Ph.D. degree in microscale magnetic components from Heriot-Watt University, Edinburgh, in 2002, 2003, and 2007, respectively. He is currently a Professor of smart systems at Heriot-Watt University. He is the Founder of the Smart Systems Group (SSG), Heriot-Watt University. He teaches smart system integration, electrical engineering, and energy systems. The research of the SSG involves multidisciplinary expertise across energy systems, sensor technologies, data analysis, and systems engineering. He is an Executive Board Member of the U.K.'s National Robotarium and the Associate Director of the U.K.'s National Centre for Energy Systems Integration. He is an IET Scholar as a recipient of the Institute of Engineering and Technology (IET) Leslie H. Paddle Prize. He is the Vice Chair of IET Scotland and an Associate Editor of IEEE Access.

JAN BERND VORSTIUS is currently a Senior Lecturer at the University of Dundee. He directs the postgraduate degree courses in industrial engineering at the University of Dundee. His research interests include revolve around tissue engineering, looking into enhancing artificial musculoskeletal tissue growth through the means of engineering approaches, thereby enabling and facilitating research to combat musculoskeletal diseases without the need for animal testing; and research in industry for extensive knowledge transfer between academia and industry, based on his expertise and interest in advanced control engineering design combined with modern digital engineering technologies.