DOCTOR OF PHILOSOPHY

Image Analysis using Deep Learning for Electrical Overhead Line Tower Management

Odo, Anicetus

Award date:
2022

Image Analysis using Deep Learning for Electrical Overhead Line Tower Management

Anicetus Ikechukwu Odo

Supervisors: Prof. Stephen Mckenna

Dr. Jan Vorstius

School of Science and Engineering

A dissertation submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

March 2022

Declaration of Authorship

Candidate's Declaration

I, Anicetus Ikechukwu Odo, declare that I am the author of this thesis. I have done all the work recorded in this thesis and cited the work of others. The contents have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

Signed: _____

Date: 24 October 2021

Supervisors' Declaration

I declare that I am a supervisor of the candidate, and that the conditions of the relevant Ordinance and Regulations have been fulfilled.

Prof Stephen J. McKenna

Signed: _____

Date: 24 October 2021

Dr Jan B. Vorstius

Signed: _____

Date: 24 October 2021

Acknowledgements

First, I would like to thank the Tertiary Education Trust Fund (TETFUND) Nigeria for the grant towards my PhD. I am grateful to the Enugu State University of Science and Technology for the nomination and financial support.

I would like to thank my supervisors, Prof. Stephen McKenna and Dr. Jan Vorstius, for their patience and guidance. It has been an eventful journey and one full of good memories. The successes recorded in this thesis would have been impossible without their supervision and timely feedback. I acknowledge Prof. David Flynn of Heriot-Watt University for his interest in my research. With his knowledge of the industry, David interfaced me with the Northern Powergrid who provided data for this research.

The outbreak of COVID-19 introduced new challenges especially with remote connectivity and supervision. I would like to thank Mahamadou Nikate, who was extremely helpful in providing the technical support needed at that difficult period of my research. I am grateful to all the members of Computer Vision and Image Processing (CVIP) Group in Computing, University of Dundee, for providing an all-inclusive learning environment. I enjoyed the weekly paper discussions and summer CVIP workshops. These events really helped to shape my research. Specifically, I would like to thank my colleagues, Mohammad Ghouse Syed and Jacob Carse for their positive comments on my work.

I would like to thank my family and friends for the opportunity to undertake this study. I thank you all for your prayers and financial commitments. Importantly, I would like to

emphasize the support of my wife, Dr. Chinasa Odo and my children, Chidimma, Chisimdi and Kamsi.

List of Publications

1. Odo, A., McKenna, S., Flynn, D., and Vorstius, J. (2021). Aerial Image Analysis Using Deep Learning for Electrical Overhead Line Network Asset Management. IEEE Access, 9:146281-146295, DOI: 10.1109/ACCESS.2021.3123158.

2. Odo, A., McKenna, S., Flynn, D., and Vorstius, J. (2020). Towards the Automatic Visual Monitoring of Electricity Pylons from Aerial Images. In proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. pp. 566-573.

3. Odo, A., McKenna, S., Flynn, D., and Vorstius, J. (2022). Identifying Paintwork Deterioration for Image-based Monitoring of Distribution Towers. IEEE Power & Energy Society (PES) Conference (Under review).

External Presentation

1. Odo, A. (2018). Towards Automatic Visual Monitoring of Electricity Pylons from Unmanned Aerial Vehicles. 2nd IEEE UK & Ireland Young Professionals Postgraduate STEM Research Symposium, University of Chester.

Internal Presentations

1. Odo, A., McKenna, S., and Vorstius, J. (2021). Electricity Tower Condition Classification from Aerial Images using Deep Learning. Computing PhD Symposium, University of Dundee.

2. Odo, A. (2019). Towards Automatic Visual Monitoring of Electricity Pylons from Unmanned Aerial Vehicle. Computing PhD Symposium, University of Dundee.

3. Odo, A., McKenna, S., Vorstius, J., Pakleppa, M., and Gkikopoulos., T. (2018). Cognitive Asset Monitoring with Health State Classification. 2nd Annual PhD Symposium, School of Science and Engineering, University of Dundee.

4. Odo, A. (2018). Cognitive Asset Monitoring with Health State Classification. Computer Vision and Image Processing (CVIP) Summer Workshop, University of Dundee.

Abstract

Electricity networks are critical national infrastructure throughout the world, delivering vital energy services and supporting interdependent assets such as transport and hospitals. The network corridors are inspected and refurnished regularly to remain useful. Due to the high number of components and the geographical spread of electricity networks, operating costs can be very high. Inspection parameters include vegetation encroachment, sagging lines, tower paintwork defects and numerous components on the towers. There are over fifty inspection parameters that are critical along a network segment. The state-of-the-art inspection process involves aerial surveys. Images are acquired and analysed manually, which adds to the high cost of aerial surveys. In addition to being costly, this process can suffer from inter-observer and intra-observer variability.

The condition-based risk management model is a popular network asset management model within the industry. The model allows for individual components rating and then enables a collective economic impact analysis for the medium and long term. The assessment model has two main stages that inform if a tower requires intervention. An aim of the routine analysis stage is to rate towers, select "at-risk" candidates for detailed investigation and refurbishment. The effectiveness of the assessment model would depend on how quick and accurate the routine assessment is able to highlight areas of need considering the limited resources during the inspection window.

This thesis focuses on automating tasks within the routine analysis stage involving image analysis. Specifically, this thesis identified towers that are at-risk of different failure modes

using deep learning. The first step in the proposed pipeline involves the identification of tower types as suspension and tension. We found that towers could be categorised automatically by focusing on the configurations at tower cross-arms. In addition to tower type detection, identifying images of cross-arm, body and foot would serve as precursor for effective extraction of specific inspection parameters. Components such as anti-climbing devices are found around tower body and not higher up the cross-arm or peak of the tower. Hence, classifying tower images to reflect the region of interest would provide a filter for object detection.

Tower conditions are often associated with the failure modes of components instances they support. This thesis demonstrated an automated detection of insulators and U-bolts as exemplar inspection parameters. Our method classified at-risk towers based on the detected instances but without explicitly labelling the instances. Instead, learning was supervised using only the condition labels of towers in their entirety. This enables us to use a real-world industry dataset without the requirement of fine-grained data annotation of thousands of individual components. While existing studies classified component instances on a tower, we classified the tower as a whole and show that tower labels are adequate for the task. Automated detection and analysis of U-bolts have not been previously reported.

Furthermore, the thesis presents the identification of paintwork deterioration for image-based tower management. We argue that the classification of tower parts may be costly. While identifying towers with immediate need for intervention, our approach to tower paintwork classification could be used as an early warning system for assets approaching end-of-life.

The utility of each sub-system has been demonstrated using a real-world industry dataset of over $7k$ towers and $300k$ images that are representative of asset failure modes and inspection scenarios.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

AP	Average Precision
AUC	Area Under the Curve
CNN	Convolutional Neural Network
CR	Condition Rating
DNO	Distribution Network Operator
DOD	Danger of Death
FP	False Positive
IOU	Intersection Over Union
LiDAR	Light Detection and Ranging
MIL	Multiple Instance Learning
NPG	Northern Powergrid
OHL	Overhead Line
OOM	Out of memory error

ROC Receiver Operating Characteristic Curve

TP True Positive

UAV Unmanned Aerial Vehicle

Chapter 1

Introduction

1.1 Research Motivation

The generation, transmission and distribution of electricity are critical energy services throughout the globe and underpin other vital services such as telecommunications, water services, transport, and education. These services require substantial investment in physical assets including overhead line (OHL) networks. Most power generating plants are located far from load stations. Networks between generation and load points are known as transmission lines. Transmission lines are usually high voltage lines. The high voltage is stepped down and made available to end-users through a network of distribution lines. To ensure continuity of services that depend on electrical power, distribution network operators inspect the assets regularly for failures or conditions leading to faults. There are several reasons for inspecting overhead line assets and these include safety and economic reasons.

The economic and societal implications of interrupting electrical supply have been demonstrated throughout the world [32]. Specific accounts relating to failure of electrical networks include blackouts in North America [4] and Europe [106]. Schmidthaler and Reichl [106] assessed the impact of blackouts in Italy and estimated the damage to society in excess of 1.15 billion Euros. The root causes of such power failures are diverse: loss of energy

generation, switching failures, IT failure as well as failures on overhead line network assets. OHL network failure has been attributed to disastrous environmental events such as the California fires that destroyed over 150,000 acres and killed 85 people [94].

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 [76, 116]. 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 the electrical networks in the United States is expected to incur a $2 trillion cost by 2030, which excludes generation asset investment [16]. 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 [75]. Miguelañez-Martin and Flynn [75] highlighted the transition from time-based maintenance (TBM) to condition-based maintenance (CBM) within the energy sector. Whereas TBM is a preventive maintenance determined by failure time analysis, CBM recommends maintenance from actual condition monitoring information [1]. A set of equipment may have the same useful life but fail at separate times because of the different environmental factors and usage. Eyre-Walker et al. [20] described an example of CBM operated by DNOs in the UK as key components of inspection and refurbishment of OHL assets.

Electrical networks have benefited significantly from the roll out of advanced monitoring technologies providing better knowledge of the electrical network [43, 72, 97]. The improved monitoring of electrical networks has created challenges and opportunities with respect to big data analytics [92]. The difficulties in managing the operational requirements and of dealing with the vast amount of data, have resulted in significant research into Artificial Intelligence (AI) to support automated decision making, control, and operational decision support for network operators [5, 15, 77].

To improve the quality of inspection and reduce costs and risks of failure, current efforts focus on remote sensing techniques such as optical satellite images, optical aerial images, and airborne laser scanners [73]. To carry out fine-grained inspection works, the focus reverts to low altitude imagery. The use of helicopters for low altitude sensing of electricity network assets is commonplace in the industry. Recently, there is an increasing usage of drone technology for short-term, single mission focus but increased resolution image capture [23]. A drone can be flown close to an asset and take high-resolution images from different views. Other alternatives to drones in terms of flexibility and cost are climbing robots [128]. Although there are proposals for the deployment of such alternatives for OHL asset inspection, they are yet to be adopted as business-as-usual capabilities. The deployment of drones, does represent an operational expense and risk [86, 108, 129, 137]. Therefore, there is a need to optimise the value of the captured images from these systems to inform predictive asset management operations. Irrespective of whether information is captured via helicopter or drone, the acquired aerial images are manually analysed by experts for faults or precursor signatures of faults. Manual analysis of towers from aerial images is a highly labour-intensive and subjective process that is error prone and expensive [20].

1.2 Overhead Line Assets Inspection Methodology

An aim of OHL assets inspection is to ascertain the current operational status or health index. The inspection procedure entails rating the condition of individual towers, which combines with previous knowledge to determine the health index of a line segment. The health index determines the network segment that must be scheduled for further detailed assessment or refurbishment during a maintenance window. DNOs analyse aerial images to determine tower ratings. As described in [20], the process includes routine assessment, which is where the bulk of the data are generated. It consists of two steps:

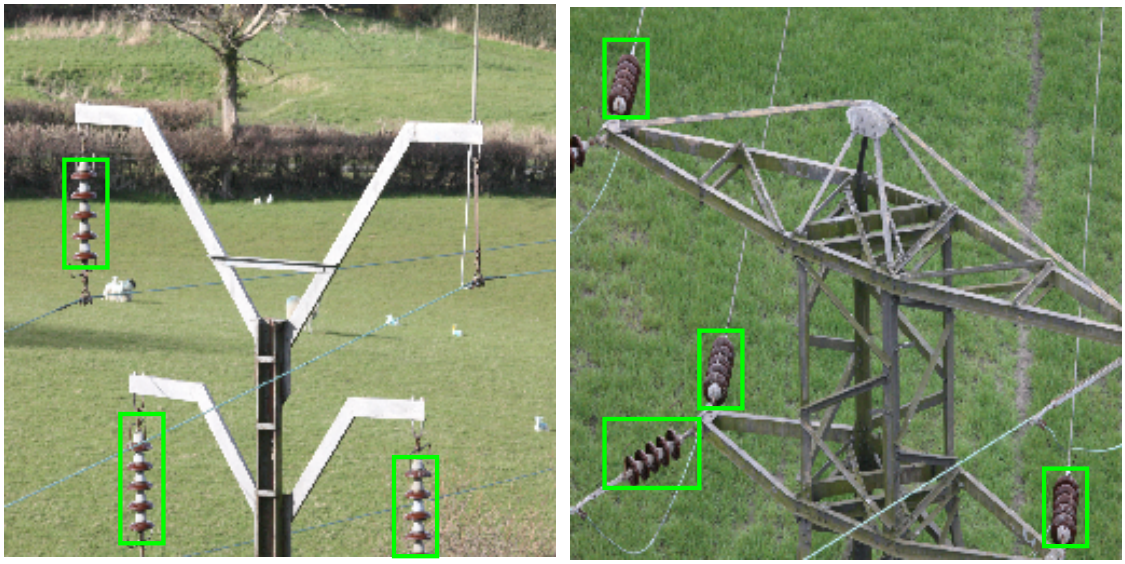
1. **Data collection**. Routine assessment of the assets is carried out from helicopters with two experts on-board (to take digital photographs and to observe and highlight (annotate) asset conditions).

2. **Data analysis**. The data collected comprising of aerial images and initial condition ratings are further analysed (off-line analysis) for purposes of making intervention decisions.

The condition ratings of each asset determined both during the survey and by further analysis become inputs for estimating the proximity to end-of-life and the consequences of failure of the asset. The use of human experts for both data collection and analysis are challenging, and results are subjective. The need arises for automating tasks along this pipeline for faster, cheaper, and more reproducible results.

1.3 Overview of Proposed Inspection Pipeline

This thesis focuses on automating the image analysis step for tower condition assessment. Corrosion is a major threat to metal structures leading to the wearing of components. Tower components such as insulators and U-bolts are prone to rust. As shown in Fig. 1.1 and Fig. 1.2, these components can be detected on two types of towers (suspension or tension). During an inspection process, detection of tower types facilitates component detection. The determination of tower types is a precursor for extracting components. To classify towers based on the components they support involves the extraction of assets on towers. The analysis of extracted components leads to the determination of tower conditions. Specifically, we aim to classify towers as healthy or unhealthy based on a collection of the assets they support. This is achieved using a combination of deep learning methods (Figure 1.3). The first stage in our pipeline involves object instance detection. This subsystem is responsible for identifying sub-image regions of insulators and U-bolts from aerial images. The second



(a) Suspension tower (b) Tension tower

Fig. 1.1 Bounding boxes showing insulator regions on towers.

(a) Insulator U-bolt (b) Earth wire U-bolt

Fig. 1.2 Bounding boxes showing U-bolt regions on towers.

stage involves tower condition classification. Multiple instances of the extracted objects from multiple images of a tower are analysed. We also automate the classification of tower paintwork failure modes.

Our methods are developed and validated using real-world electricity network inspection data set comprising of visual images that are representative of the diverse component types and failure modes encountered in real-world inspection scenarios.

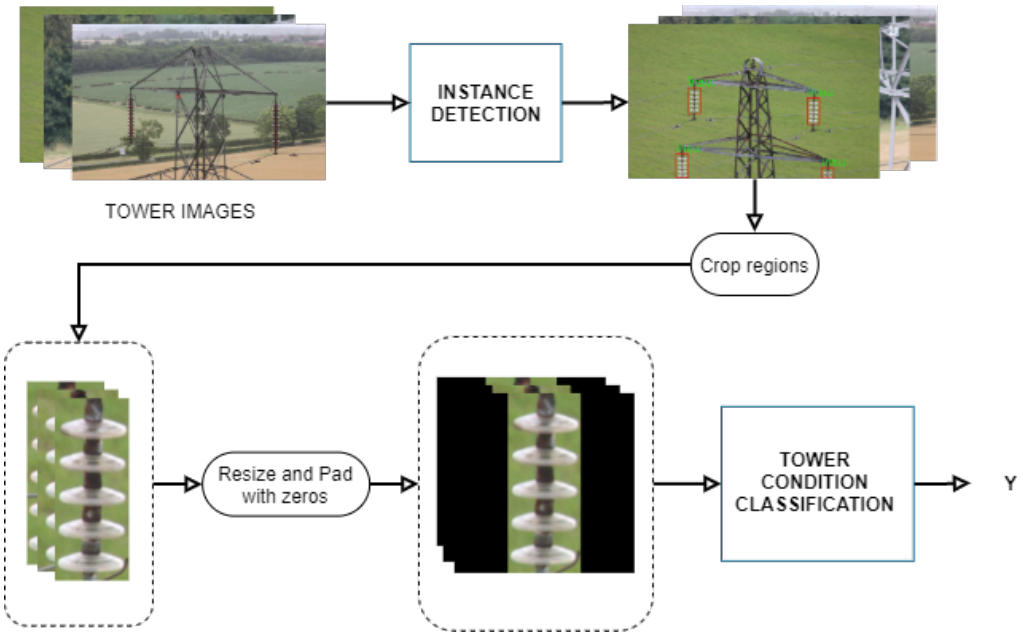


Fig. 1.3 Overview of inspection pipeline for overhead line towers with two stages: (1) Instance detection to identify the regions of interest, (2) Tower condition classification to rate towers based on the detected instances.

We explored two methods for the classification of tower conditions namely instance classification and multiple instance learning (MIL). In the instance classification technique, all detected object instances on a tower are classified and a tower label is determined by aggregating instance class predictions. The MIL-based technique classifies sub-bags of instances and aggregates sub-bag class predictions to determine tower labels.

1.4 Research Contributions

The main contributions of the thesis are:

1. Detection of four insulator types simultaneously, i.e., porcelain, ceramic, glass, and composite.

2. Detection of U-bolts from aerial images. This thesis includes the first published study on automated detection of U-bolts from aerial images [82].

3. Classification of towers' configurations as suspension or tension, which is a precursor for the detection of assets on a tower [81].

4. Introduction of tower condition classification. While existing studies try to classify every instance of detected components on a tower, this thesis classified towers. Our approach achieved tower classification without explicitly labelling instances of insulators or U-bolts. It demonstrates that tower level labels are adequate for the task thereby reducing the requirement of data annotation [82].

5. Classification of tower paintwork directly from whole tower images instead of estimating the defect on different parts.

6. Evaluation of the proposed assessment pipeline using a real-world inspection dataset, which is representative of the diverse component types, scenes, and views.

Minor contributions include:

- Comparative evaluation of Mask R-CNN and RetinaNet for electricity network assets identification. To the best of our knowledge, these models have not been previously comparatively analysed for this application.

- Application of multiple instance learning for the analysis of electrical OHL assets.

The focus of this thesis is not the development of new deep learning methods, per se. Rather, it is interested in the development of novel techniques, combining existing methods to solve difficult inspection applications.

1.5 Structure of Thesis

The rest of the thesis is structured as follows.

- *Chapter 2: Literature review.* This Chapter reviews the current state-of-the-art on electricity overhead line tower assessment from aerial images. The focus of the chapter includes a review of methods and the potential of deep learning for automated assessment of towers.

- *Chapter 3* describes the real-world data set used in this thesis. It also describes some design considerations including annotations of instances for developing detectors and preprocessing of the image data.

- *Chapter 4: Classification of Tower Configurations.* This chapter reports the results of tower classification as suspension or tension. It uses majority voting as an aggregation function for multiple images of each tower.

- *Chapter 5: Detection of Insulators and U-bolts on Electricity Tower Images.* This chapter reports on automated extraction of assets from tower images. It investigated the effect of occlusion and size of objects on the detection of insulators from aerial images. It reports on the importance of multi-tasking for object detection.

- *Chapter 6: Tower Condition Classification.* This chapter describes the various methods considered for this task including instance and multiple instance learning. It investigated multiple instance learning when training with resource constraints. The methods in Chapters 5 & 6 were evaluated using real-world inspection data. Insulators and U-bolts are considered as exemplars of tower condition parameters.

- *Chapter 7: Classification of Tower Paintwork Failure.* This chapter reports automated assessment of electricity overhead towers to determine the status of their paintwork. Specifically, high-risk towers are classified against the rest of the towers that are within normal operating conditions.

- *Chapter 8: Automated Tower Assessment System.* The different components of the proposed routing assessment pipeline are combined into a system. This represents a use case scenario and shows the possible input and output requirements.

- *Chapter 9: Conclusions.* This chapter presents a summary of the main findings of this thesis, describes the limitations of the experiments, and proposes future directions of research in this area.

Chapter 2

Literature Review

2.1 Introduction

Electrical overhead lines comprise of components that are connected in a special way to deliver energy to point of use. These assets are integral to the profitable delivery of the services that businesses depend upon [14]. The tangible assets on a distribution network include transformers, high voltage towers and the components they support such as conductors, insulators, and U-bolts. Various environmental factors such as tree encroachment and bush fire along power lines need to be monitored to avoid power cuts. The California fires in 2019 caused a wildfire in Northern California that destroyed over 150k acres and killed 85 people [94]. The incident in Texas was caused by extreme weather events, disrupting renewable energy generation, and leading to overload of the distribution network [52]. Several other cases of blackouts have been traced to these events and the economic and societal implications are enormous. Other accounts relating to failure of electrical networks include the blackouts in North America [4]. Schmidthaler and Reichl [106] assessed the impact of power failure in Italy and estimated the damage to society to the tune of 1.15 billion Euros. The root causes of such power failures are diverse. The degradation of network assets has also been linked to disastrous environmental events such as the fire incident in California [94]. Component

failures and human errors were reported to be responsible for a third of blackouts world-wide between 2011 and 2019 [32].

As a result of the multifaceted implications of blackouts, operators in this industry are required by law to inspect their networks. In the UK for example, section 24 of the Electricity Supply Regulation 1988 requires distribution network operators (DNOs) to take practical power line inspection steps to avoid unplanned outages [49]. Therefore, regulators of the energy industry have a well-developed set of standards for assessing network assets. Network operating costs for distribution network operators in the UK amount to over £6.31 billion between 2018 - 2022, covering faults, tree cutting, inspection and maintenance programmes [83].

Inspection of OHL assets is a challenging and costly process. Inspection service is expected to go on without interrupting critical services. As such, OHL surveys are scheduled to minimise down-time. The use of men on patrol for monitoring OHL corridor and tower components is now out of fashion because of the high risk involved. Men on patrol are inefficient considering the coverage of the network and in terms of response time. The current state-of- the-art for OHL inspection deploy from aerial platforms using visual and lidar systems. Non-intrusive surveys of electrical assets are safe and economical and platforms for delivering aerial photography are well developed. Helicopter survey is the current industry standard for aerial inspection but more recently, unmanned aerial vehicles (UAV) are becoming popular. UAVs can provide images with increased resolution [23], which are better for automated image analysis. In addition, they are safer and easily maneuverable compared to helicopters.

2.2 Overview of Existing Methods

Current OHL assets inspection deploys aerial sensors (visual or lidar) to obtain non-intrusive survey that is both safer and more efficient than manual assessment. However, the acquired

images are manually inspected, a process that is highly labour intensive and prone to subjective human interpretation. To improve on the current pipeline, research into automating tower image analysis is encouraged. Methods commonly reported for analysing OHL assets images are predominantly domiciled within the computer vision and deep learning (DL) with dominance of DL-based solutions.

Deep learning is a subset of artificial neural networks and provides promising approaches for automated decision support systems. Convolutional neural networks (CNN) are important for learning representations from images. CNNs have been used for detection, classification, and segmentation tasks with high success in many domains. Depending on the complexity of task, CNNs can be scaled to extract better features by increasing their depth and width. Deep networks are useful and robust for some tasks, but there are some challenges in training these models. They typically require at least thousands of correctly labelled data points, depending on task, to achieve good results. There are different approaches to learning and these can be classified under two broad types, namely supervised and unsupervised. Fully supervised learning requires a label for every data point which can be costly to obtain. Semi-supervised settings enable the use of labelled datasets in combination with, typically much larger, unlabelled datasets. In a supervised learning scenario, the process of re-using and adapting a network with previously learned features is known as transfer learning. The scarcity of correctly labelled data necessitates alternative training pipelines in what is known as weakly supervised learning [118]. One form of weakly supervised learning is multiple instance learning (MIL) in which each data sample is presented as a bag of different instances.

The successes of DL networks have resulted in state-of-the-art applications e.g., medical image analysis [10]. Recently, EfficientNet, a method for analytically determining the depth, width, and resolution of convolutional neural networks (CNN) for optimised performance was introduced [114]. EfficientNet achieved state-of-the-art results on ImageNet classification challenge. There are many deep networks used for image classification tasks and have

been used as backbones to extract rich features from images. ResNets are CNNs for image classification with micro networks within the network. The micro networks are called residual modules and connecting multiple residual modules facilitates the training of deeper CNNs. Each residual module has two branches: a stack of convolutional block and the residual link (input). Kaiming et al. [53] used bottle-neck structure building blocks. Each residual module is a system of Convolution-Batch-Normalisation-Activation layers. Activation before convolution within residual modules was found to improve classification accuracy in later versions of ResNets [36].

Object detection involves the localisation and classification of sub-image regions. Girshick et al. [27] introduced a region-based object detection network (R-CNN), which combines CNNs for localisation and classification. A series of region-based detectors evolved from the original model including Fast R-CNN [26] and later, Faster R-CNN [96]. The later method found that sharing features between the ResNet backbone and the detection arm of the network improved the speed and achieved 5 frames per second [96]. Mask R-CNN [35] is an improvement on Faster R-CNN, which in addition to detection, provides instance segmentation. Two-stage networks achieve good detections but at high computational costs compared to one-stage detectors.

RetinaNet is a good example of one-stage model designed to bridge the gap between one-stage and two-stage networks by eliminating the bottleneck posed by background-foreground class imbalance [59]. It was based on ResNet architecture and feature pyramid backbone. Other examples of detection using a single deep neural network include the single-shot multi-box object detector (SSD) [64] and "You Only Look Once" (YOLO) [95]. The original SSD used a VGG-16 [112] as backbone. There are several versions of YOLO up to the fourth generation, with successive model improving on its predecessor.

The effectiveness of the detection networks listed above have been tested on a standard public data set considering speed and accuracy [59]. While two-stage networks have better

precision, one-stage models are faster. RetinaNet has been limited to the detection of pins from overhead tower images [121].

2.3 Sensing Platforms

Several techniques have been tested for surveying the power line corridor. These include the use of satellite, deployment from low altitude aircraft like helicopters and more recently, unmanned aerial vehicles (UAV). These platforms are usually equipped with sensors such as radar, thermal sensors, LiDAR, and visual devices [73]. Currently, low altitude aircraft such as helicopters mounted with visual sensors (digital cameras) are used. Towers are scheduled for inspection and flights are carried out with pilot and crew. During each flight, a crew of specialists in the field (engineers) come aboard for online assessment of assets. Images are collected and condition ratings were recorded during flight. Secondary off-line assessments are carried out for selected network routes. The process of data collection from helicopter is labour intensive and costly. An alternative inspection platform based on unmanned aerial vehicles (UAVs) is becoming popular. UAVs have special features, which make them suitable for aerial surveys. A helicopter is bigger, has people on-board and weighs more than a UAV. In contrast, UAVs are remotely flown and are flexible to manoeuvre for data collection [37]. UAVs are low-cost equipment compared to helicopters and other platforms [8]. Additionally, UAVs can be deployed very close to the asset without endangering the system or the network [58] and can capture data good enough for component level inspection.

We could not access UAV-based inspection data for our experiments. In this thesis, we report a real-world inspection data set collected from helicopters. This data set is the closest we could find at the start of the research. The images provided will provide information regarding the characteristics of aerial images of electrical networks and most importantly, how the various image attributes affect the assessment and management of OHL assets. Although UAVs are gaining popularity in the research community, for surveying power line

corridors, it is yet to be accepted as business as usual by the industry due to issues of flight safety and data protection. The public perceives UAV as infringing on privacy because they are remotely operated and can collect business as well as personal information, sometimes without notice.

The safety of flight along a power line corridor is key to a successful inspection be it from manned or unmanned aerial platforms. There is a high possibility of accidents due to system, human or environmental factors. Environmental factors that affect the navigation of aerial platforms include gust wind and poor illumination. Liu et al. [62] proposed the creation of a no-fly zone along the distribution network corridor using GPS coordinates of the towers to solve the problem of gust wind. Due to the lightweight of UAVs, they could easily go adrift by wind and collide with power lines.

Researchers have proposed several tracking methods for safe navigation of the networks. Sa et al. [101] investigated vertical take-off and landing for the inspection of pole-like structures. They combined monocular, inertia, and sonar data to help with navigating the aircraft. Kalman filter has been employed by several authors to detect and follow applications. For example, [101] employed the extended Kalman filter to maintain a safe distance from a pole structure even in the presence of environmental disturbances. Golightly and Jones [28] also combined Hough transform and Kalman filters to guide a rotor-craft along detected power lines. Jones et al. [50] used an air vehicle simulator (AVS) to demonstrate that visual data can be used to determine and regulate vehicle position relative to the overhead lines. Cerón et al. [11] developed a system to detect and follow power lines from images. It is evident from these studies that visual data can provide enough information for flight path planning, which is an important step in the automation of power line inspection. Authors in [132, 135, 136] have proposed different techniques for estimating the position of the assets and landmarks in relation to the UAV.

2.4 Overhead Line Assets and Failure Modes

Electrical overhead line networks comprise assets for transmission and distribution of energy including towers and tower fittings. Fig. 2.1 shows a tower and several fittings like insulators and cables. Electricity networks span communities, cities, and vegetation. While surveying the lines, other components are sensed. The detection of obstacles like vegetation and buildings along the power line is a useful aspect of automated assets inspection. The power line is usually a corridor, which ideally should be free of obstacles like trees. However, the line cannot be completely free from encroachment.

Tree cutting takes a huge chunk of the investments on OHL network maintenance. Automatic measurement of vegetation in breach of the power line corridors from visual images has been reported [51, 98]. There are other events on towers and lines that facilitate the degradation of network components and should be detected. For example, Towers are inspected for activities of birds such as bird nests and accumulation of droppings [13, 34, 93]. Bird droppings on insulators cause flashovers and power cuts. Zhai et al. [131] investigated icing accumulation and employed Canny, Sobel, and Adaptive Weighted Sobel techniques. They show that Adaptive Weighted Sobel performed better than the other edge detectors. Due to the danger of prolonged icing on electrical towers, several authors have investigated this event [30, 31, 55, 40, 41]. Early identification of these natural events is important as that could facilitate quick intervention programs and avoid network failure.

2.4.1 Detection of towers

The detection of tower is perhaps the first task for aerial surveys of OHL networks. Most of the components inspected are on the tower. The tower supports cables or lines and other components like insulators and U-bolts. A number of researchers have investigated electrical pylons from images [19, 102]. Dutta et al. [19] used gradient and cluster density-based segmentation to minimise the clutter arising from heterogeneous background before

detecting pylons and found that the method could detect pylons with the absence of false positives. Detection of tower lattice is a useful step for the identification of their condition parameters. There are different pylon designs and automatically categorising these may lead to quick assessment of the components they support. Sampedro et al. [102] proposed an automated pipeline to detect tower structures in images and classify them using multi-layer perceptron. Histogram of oriented gradients (HOG) features were used to discriminate between foreground and background in a sliding window fashion to suggest candidate tower regions. The selected candidate was classified as one of four tower structures. In addition to tower detection, it will be useful to identify their functions on a network segment, e.g., a tower can be used as either suspension or tension.

2.4.2 Detection of Insulators

The condition ratings of towers on a line segment are usually determined by the aggregate status of the components they support. Therefore, there is a need for fine-grained detection of assets like insulators from towers. According to a recent review [65], insulators and insulator failure modes are the most studied of OHL assets. Insulators can be described as devices used in fitting a cable to the tower as shown in green boxes of Fig. 2.1. Insulators have been detected in literature [46, 66, 80, 130]. Jabid and Ahsan [46] proposed encoding insulator features using rotation invariant local directional pattern. These features were then learned by an SVM to classify regions of the insulator. There are 2 categories of faults assessed on insulators such as their electrical and mechanical failures. Insulator electrical failure modes are faults that reduce or undermine the electrical properties of the device and could lead to energy drain. Examples of insulator electrical failure modes are crack or chip insulator sheds and sometimes completely missing sheds. Cracks and chips on sheds are typically found on porcelain insulator types. Glass insulators would shatter. Insulator mechanical rust has to do

with rust or corrosion on the insulator string. Insulator mechanical rust could be local around the cap region or the pin and in some cases affecting the entire device.

While there is a handful of reports handling insulator electrical type of failures [21, 24, 42, 61, 78, 111, 115, 123, 125], insulator mechanical failure is not as popular. Jalil et al. [47] reported the detection of insulator instances as either rusty or normal. It is useful to classify every instance of the multiple insulators on each tower. However, it may be too costly to label every insulator instance, which may be the reason the industry handles failure modes differently. Instead of detecting and classifying individual instances, the industry rates a tower, albeit manually.

Computer vision traditionally relies heavily on feature extraction (e.g., shape and colour) of assets and present fast processing. Specifically, insulators were detected in [80, 130] using their shape attributes. Potnuru and Bhima [90], Zhang and Yang [133], Fang et al. [21], Xin et al. [123] explored colour space transformation for insulator detection. To enhance performance, [39] explored the fusion of features such as gray level, colour name, and histogram of oriented gradients (HOG) features for insulator detection. Insulators come in different shapes and colours. Shape and colour may not be sufficient in situations where the object has poor contrast with the background, i.e., glass insulators. Yu et al. [127] highlighted the difficulty of segmenting glass insulators in low contrast images and combined texture and shape features for the task. Thus, texture, shape, and colour are important features for detecting the insulators in images. An instrument based on such hand-crafted features may be difficult to generalise because insulators are so diverse [61]. Therefore, a model that can learn these features automatically would be preferred.

Deep learning bypasses the requirement to manually define discriminating features. Within the OHL assets image analysis, several convolutional neural networks (CNN) have been reported. CNNs used for object detection can be categorised into one-stage and two-stage networks. A two-stage network incorporates two CNNs: the first for region proposal

and the second to localise and classify the object of interest. Examples of two-stage detectors are Faster R-CNN [96] and Mask R-CNN [35]. Both networks have achieved a state-of-the-art results in object detection [24, 58, 61]. On the other hand, a one-stage detector achieves object detection using a single network.

Gao et al. [24] used a two-stage model for the extraction and analysis of insulators. The first component is the region proposal network and employed a VGG16-based network to suggest regions in the image with insulator instances. Faster R-CNN was also applied in [24, 61] for the detection of faulty insulator caps. While [61] explored glass insulators, [24] worked on porcelain and composite types. Global detection and local segmentation of glass and ceramic insulators were presented in [56] and they employed 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. A system for the detection of insulators was proposed by [103] in which a fully convolutional network (CNN) derived from a modified U-net was employed. The network was trained within a generative adversarial network framework with some transfer learning. Li et al. [56], Sampedro et al. [103] detect insulators and highlighted missing sub-insulator regions. This requires annotating each insulator sub-region and is costly considering the vast number of components on an OHL network.

Nguyen et al. [79] highlighted speed as a major consideration for using one-stage networks. Examples of one-stage models are Single-shot multi-box object detection, SSD [64], YOLO [95] and RetinaNet [59]. These models have been cited in the literature for OHL assets detection [12, 74, 79, 84, 121]. The method in [74] included fine-tuning an SSD-COCO model with some aerial images, which they called basic data and comprising of different insulators and backgrounds to derive a basic insulator model. 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. Another one-stage network named YOLO was employed in

[12, 84] in which insulators have been detected. In [33], a ResNet was applied to extract features of insulators for a YOLO and use different input scales. Chen et al. [12] enhanced blurred images using Super-Resolution Convolutional Neural Network (SRCNN). While [24, 74] detected porcelain and composite insulators, [61] was interested in glass insulator type. However, [84] noted that a tower could be supported with a combination of insulator strings. It would be more efficient to detect the different insulator types simultaneously.

The results of the reviewed papers and specifically [56, 103] are compelling in terms of performance and they are related to the current study on insulator detection. However, instead of instance-level condition assessment, an alternative approach is to classify the tower as a whole. Automatic tower classification is a step further as it does not require exhaustive annotation of individual components.

2.4.3 Detection of U-bolts

The detection of U-bolts is required since insulator U-bolts bear the load of the circuit cable they support. This component may not have been painted and would rust at the link point and then fail. This critical tower condition parameter has not been reported in the literature. In addition to component detection, this thesis will focus on the automatic classification of towers based on the detected fitting such as insulators or U-bolts.

2.4.4 Detection of lines (Conductor)

The conductor has been investigated by many researchers like [57, 109, 117, 124]. Yang et al. [124] converted video frames into binary images using adaptive thresholding and detected line candidates in the binary images by Hough Transforms. This was followed by clustering using Fuzzy C-means (FCM) algorithm to discriminate between the power lines and other line candidates. The technique showed successful discrimination of power lines from patterns like roads, the edge of a river, and vegetation. A similar study in [57] used

Pulse Coupled Neural Network (PCNN) filter to remove background noise from images before Hough transform was employed to detect straight lines. Thereafter, knowledge-based line clustering was applied to refine the detection results. The experiment showed that applying knowledge to Hough line detection significantly improved detection from UAV-based image data. Bhujade et al. [7], Sharma et al. [109] suppressed the natural surroundings (region of sky and vegetation) and used morphological erosion operations that suppressed patches in the thresholded binary image. This was followed by Hough transform. In [117], power lines were extracted based on directional constraints using double-side filter. The improved Hough transform with parallel constraint was used for power line recognition. The results show significant improvement in power line recognition because of the addition of direction & parallel constraints. Zhu et al. [139] presented a double-side filter-based power line recognition method for UAV vision system. Their method was based on linear object enhancement and parallel lines constraint just as in [117], and radon transform was used to find the parallel lines that characterised power lines.

Still, visual images have been used in many studies for the investigation of power line detection. However, [67] presented a real-time power line extraction from video signal. Others like [45, 105] have shown the extraction of power lines from 3D LiDAR scanning. Ippolito et al. [45] utilised a voxel-based method with a series of classifiers to identify and reconstruct power lines. A mini-UAV mounted with LiDAR (AOEagle) was proposed in [105] for sensing the power line corridor, for the detection of poles, transmission lines, span, and sag. Most of the reviewed studies on the detection of cables focused on detecting the conductor span. Fangzheng Zhang et al. [22] found abnormalities/defects such as broken cable strands in images of transmission lines. They achieved this by first eliminating the effect of fog and using histogram specification to detect the lines.

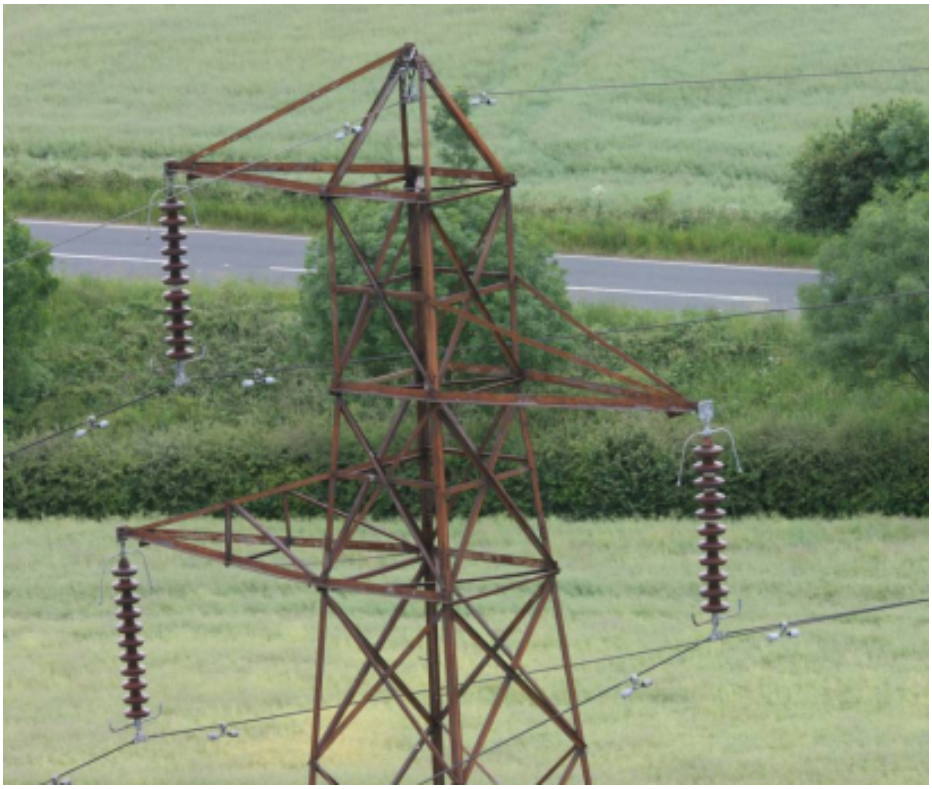


Fig. 2.1 Tower at high risk of paintwork deterioration.

2.4.5 Tower paintwork degradation

Electrical overhead line towers are built from different materials, e.g., wooden and metal towers. While wooden poles rot, especially at the foundation (footing), metallic towers corrode (Figure 2.1). Early detection of these conditions is an important inspection requirement. The prevention of corrosion using paint coatings is a common practice in many applications including bridges, marine vessels, pipelines, and electricity towers. Automatic visual inspection has become a popular alternative to the manual estimation of paintwork defects across different domains.

Computer vision and machine learning techniques have been applied for the automation of image analysis of electricity towers [19, 71, 81, 82, 102, 130]. A method for estimating corrosion from outdoor telescopic images of electric towers was presented by [119] and used a support vector machine with radial basis functions. Synthetic images were added to originally

acquired data to augment for training. Colour information is useful for detecting paintwork defects on metal surfaces in different applications [3, 6, 38, 89, 104]. A detailed comparison of the effect of colour transformation on image classification has been reported [29]. While Hue Saturation Value (HSV) colour space performed better than RGB in [119], LAB was better for CIFAR-10 dataset [29]. This means there is no one colour fits all for image analysis. A colour space suitable for a particular task would need to be carefully determined in advance. In this paper, we experiment with RGB colour space and compute RGB mean subtraction as a normalisation technique. This technique helps to reduce the effects of lighting variations during image classification [100].

Machine learning is popular for the detection of paintwork defects in several domains. Specifically, it can provide a quick assessment of tower paintwork conditions. A support vector machine was applied for corrosion in pipes [38] and a simple perceptron for corrosion on vessel hull [85]. Both methods were based on features extraction (i.e., colour and texture). Convolutional neural networks (CNNs) are well established for object detection, segmentation, and image classification. CNNs have been used for paintwork classification [2, 6, 17, 63, 89]. CNNs often perform better than traditional computer vision methods. However, the scarcity of properly labelled data is a challenge. Where there is a lack of sufficient data, practitioners adopted transfer learning as a method of approximating a model's capability on a new task. Petricca et al. [89] fine-tuned AlexNet using 1,300 rusty images and 2,200 non-rusty images and evaluated on 100 images comprising 37 rusty and 63 non-rusty images. Zhang et al. [134] used 700 cropped corroded regions. Cropping corroded sub-image levels and annotating them can be very expensive and may explain why there is a limited number of well labelled image data for the application.

A deep extreme learning machine with local receptive field was used by [68] to estimate paintwork levels using a small number of tower images. Their private data set comprises 2,797 images of cropped tower parts at different levels of paintwork failures/corrosion. There

are three classes in their data set comprising of 940 images with newly coated steel-work (class A), class B has 1,184 images of regions with cracked and peeling paint coating, class C has 673 images of exposed and rusty metal structures. The motivation for a deep extreme learning machine was the limited amount of data. An extension of [68] was presented in [69] and combined images and text data to improve performance. While [68, 69] used parts of towers, i.e., sub-images of 50×50 pixels for learning, this thesis classifies towers without explicitly labelling tower parts. Identifying at-risk-towers using tower labels would potentially reduce the cost of data annotation and speed up processing.

2.5 Challenges of Deep Learning Approach

Although traditional computer vision tools have been used extensively for the analysis of OHL assets, deep learning methods have dominated the research in this area. The reason for the popularity is that DL learns features automatically instead of using predefined features. Deep learning methods have shown real value in domains such as medical image analysis and automated driving. However, there are some requirements of this method that are still very challenging. Deep learning requires a good quantity of properly labelled data to achieve good results. Unfortunately, the process of collecting and annotating aerial images is costly. Also, there are several levels of restrictions especially with regards to protecting company business as well as data protection guidance. Different proprietary data sets such as those mentioned above are in use by researchers to answer the same question of automating the inspection pipeline. Because of insufficient quantity, authors increase the size of small data sets by applying several augmentation protocols [33, 84, 115]. Ohta et al. [84] sourced image data from the internet.

Other challenges include an imbalance between the normal and abnormal targets. For example, in [115], 60 defective insulators were photographed from towers. To solve the problem of imbalance between these and the non-defective examples, 996 synthetic images

were generated from the 60 original images. Most of the previous works on tower image analysis e.g., [24, 33, 47, 61, 74, 84, 115] have relied on augmenting relatively small data sets to try to avoid over-fitting. For example, [47] used a data set of 132 images from which 160 insulator regions were labelled. Zhou et al. [138] reported a significant number of images for insulators and vibration hammer detection. It would be useful to know how the images were collected, i.e., the distribution of images across towers. Images from the same tower will likely present the same failure signs and therefore should be properly partitioned to avoid bias.

Jiang et al. [48] presented a combination of visual and infrared image data. Fusing multiple sensors may enhance the representations and possibly improve the performance of asset detection. Data fusion may be helpful in areas where visibility is challenging. In this thesis, we make use of a real-world electricity network data set from many years of assets inspection. The data comprises visual images and are representative of the diverse component types and failure modes encountered in real-world inspection scenarios.

For traditional computer vision methods, the features are defined in advance. In contrast, deep learning methods were designed to learn the association between features. In other words, they derive the discriminating features required to solve the task. Although deep learning has shown superior performance in many application areas over traditional computer vision methods, the question of what the network's decision is based on persists. Selvaraju et al. [107] presented a technique to visualise CNN predictions using gradient based class activation mappings (Grad-CAM). The method provides some clue about the features in an image responsible for image-based classification tasks.

2.6 Summary

The survey of electrical overhead lines for prognostics purposes is an age-long practice, traditionally, by humans on ground patrol. Due to the span of the networks and the need

for quicker response time, ground patrol is out fashioned and replaced by aerial platforms. For fine-grained component level inspection, low altitude image collection methods have dominated the inspection space with the current standard being from helicopters. Helicopters are bulky to fly especially with the hovering requirement over multiple towers. There are numerous proposals for alternative aerial platforms and the development of unmanned aerial vehicles is prominent within the research community.

There are several components inspected by distribution network operators along the power line and on the tower. The last two decades have witnessed a high volume of research focusing on the use of image processing techniques including deep learning for applications in medicine, autonomous vehicle, and importantly for the automation of OHL assets inspection. Research has demonstrated that deep learning can deliver high precision in image analysis for detection, classification, and segmentation tasks. However, the lack of properly labelled data in the right quantity is a major challenge for this method. The use of UAVs for aerial photography can lead to the quick collection of OHL assets data. DNOs are looking at the demands of the future including electric vehicle charging and smart streets. New demands mean new challenges and would need methods for network reinforcement. Digitisation and automation of OHL inspection pipelines are required for efficient energy service delivery and there is evidence in the literature that the industry is investing along that line.

Some components of electrical towers have already been detected. Some authors focus on events that could lead to faults like the activities of birds, icing, and vegetation. Most of the reviewed papers target the failure of specific components like insulators. U-bolt has not been detected. For insulator detection and condition classification, existing studies annotate sub-component faults like segmenting missing insulator sheds. Considering the huge cost associated with labelling data and the vast amount of assets/components on the network, and the volume of inspection parameters for each asset, classifying instances may not be the best solution. Within the context of condition-based risk management, the industry is not

practising instance-level assessment. Instead, the tower condition is an aggregation of the condition ratings of multiple components.

This thesis is interested in detecting electrical tower components and will investigate tower condition parameters like paintwork and faults arising from insulators and U-bolts. The identification of tower functions is fundamental to the detection of faults. For example, the classification of towers as suspension and tension defines how insulators are deployed on them. They also determine where on a power line the tower could be used.

Chapter 3

Data sets

3.1 Introduction

Electricity networks in the UK like in most countries in the world are grouped by the nature of their function into transmission and distribution networks. Electricity transmission networks transmit high-voltage electricity from the generating plant to the point of distribution networks across the country. In the UK, the transmission network is owned and maintained by National Grid Electricity System Operator. Distribution networks carry electricity from the high voltage transmission grid to industrial, commercial, and domestic users. There are 14 distribution network operators (DNOs) in the UK, owned by 6 different groups. A member of this group is the Northern Powergrid, comprising of the Northern Powergrid (Northeast) Limited and Northern Powergrid (Yorkshire) Plc.

The real-world data set presented in this thesis was provided by the Northern Powergrid. The data comprises of still high-resolution images taken using helicopter inspection survey of electricity overhead lines between 2010 - 2015 [20]. They also provided inspection reports that describe the status of the assets at the time of assessment. The inspection methods involved a cameraman and an observer both of whom are engineers with good understanding of the information required for the inspection task. Photographs were taken

using digital camera fitted with an optically stabilised telephoto lens. To capture sufficient details, multiple images were taken. For some inspection parameters like insulators and U-bolts, towers were surveyed from both left and right circuits. Each circuit was inspected separately. When standing under a line and facing the structure, the left-hand side is defined as *left circuit* and the right hand side is defined as the *right circuit*. The reason for inspecting the circuit separately might be to make identification easier during subsequent assessment and refurbishment. For example, route A136 in T7799 (Appendix A) presented inspection data for the left circuit only suggesting that the towers are single circuit. Route BCN presented in T6775 (Appendix B) have double circuit towers. The condition ratings for right and left circuit tend to agree most of the time. To combine the three databases, we merged the right and left circuits using a max operator. This implies that tower label represents the highest condition rating of either left or right circuit.

The focus was on the condition parameters such as the earth-wire, insulators, and fittings along the cross-arm region of towers. Other condition parameters surveyed include conductors, tower paintwork, tower footing as well as vegetation encroachment on the base of the tower. The job of the observer was to highlight flash points that require more attention. The flights were carried out during spring and summer months between April and October to allow for better weather and lighting conditions. On the average, 53 images were taken from each tower.

Three data sets, T6775, T7638, and T7799, were provided. Each data set consists of tower condition information and images collected from lines and towers scheduled for survey at different times. The data sets provided were surveyed in 2011, 2014 and 2015, respectively. The lines inspected in T6775 and T7638 comprise mostly of high voltage towers. T7799 has mainly low and medium voltage towers. The data sets have a total of 6,829 unique towers (i.e., 4,823, 993 and 1,013 towers in T6775, T7638 and T7799, respectively). Some towers (233) had been inspected more than once across the data sets. For example, 219 towers

Table 3.1 Tower condition parameters

Asset	Condition parameter
Tower footing	Check tower foundation for cracks on concrete Check for painting on concrete muffs
Tower body	Check for presence of anti-climbing device on tower Check for danger of death sign How many sides are with number plates? How many sides have Safety Sign?
Tower steel-work	Check paintwork failure
Cross-arm	Insulators (missing/chips insulator sheds) Insulator (rust) Insulator U-bolts or Tower attachments (rust) Earth-wire U-bolts (rust)
Conductor	Broken conductor strands Spacers Vegetation encroaching online and tower footing

presented in T6775 were scheduled again in 2014 (i.e., T7638). Also, 14 towers that were inspected in 2011 (T6775) were assessed again in 2015 (T7799). These towers were surveyed at different times and therefore are considered as separate tower inspections with a total of 7,295 towers. Condition assessment reports (Excel Spreadsheet format) were provided for each route i.e., 32 routes in T7799, 21 routes in T7638, 148 routes in T6775.

3.2 Data Extraction and Clean-up

Each tower was presented as a folder comprising of multiple images (i.e., a bag of images). A list of all the towers inspected along each route were reported. Tower names and their inspection parameters e.g., insulators, U-bolts and paintwork were extracted from the spreadsheets.

Table 3.2 Rating tower paintwork condition

Rating	Meaning
CR-1	new [Green]
CR-2	Paint - slight (less than 10%) wear, or primer visible through topcoat. [Green]
CR-3	Paint - moderate (between 10% and 50%) wear, paint cracking, some peeling. [Orange]
CR-4	Paint - severe (more that 50% wear or paint peeling, AND/OR any sign of paint blistering with evidence of severe rust underneath paint, painting required urgently. [Red]
N	Not applicable - tower has never been painted.
U	Unsure
?	Refer to network owner [Red]

Table 3.3 Criteria for rating insulator mechanical rust

Rating	Meaning
CR-1	As new [Green]
CR-2	Up to 10% rust. [Green]
CR-3	Between 10% and 50% rust. [Orange]
CR-4	More than 50% rust. [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]
M	Missing [Red]
N	Not applicable
U	Unsure
?	Refer to network owner [Red]

3.3 Design Considerations

The main condition parameters inspected on each tower are presented in Table 3.1. This thesis focuses on three of these namely tower paintwork, insulators, and insulator fittings,

Table 3.4 Criteria for rating U-bolts or Tower attachment

Rating	Meaning
CR-1	As new [Green]
CR-2	Between 10% and 50% of area affected by rust, and/or up to 10% wear [Green]
CR-3	More than 50% of area affected by rust and/or between 10% and 50% wear [Orange]
CR-4	More than 50% rust AND more than 50% wear indicative of potential failure or already broken [Red]
CR-5	Exceptional circumstances [Red]
M	Missing [Red]
N	Not applicable
U	Unsure
?	Refer to network owner [Red]

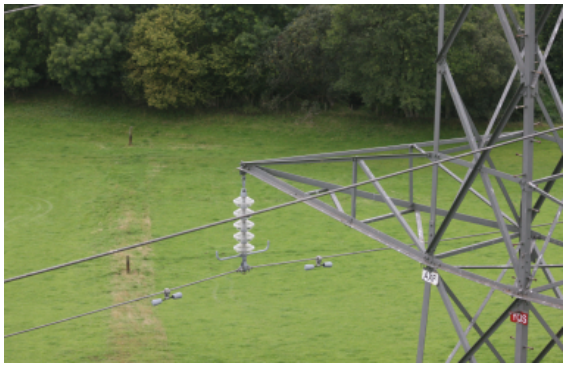
e.g., U-bolts. The agreed criteria for assigning condition ratings for the selected failure modes are shown in Table 3.2, Table 3.3 and Table 3.4, respectively.

Images in the data set were labelled to reflect the region of tower and more importantly, the target component. As shown in Table 3.5, there are about thirteen categories including Insulator Footing, Insulator Top, Insulator Middle, Middle, Footing, Top, Anti Climb Guard and Earth Wire. Others are whole tower, signage, cable platform, substation, and spacer. Examples of some image categories are shown in Fig. 3.1, Fig. 3.2 and Fig. 3.3.

The data set was considered useful for several reasons: (1) the data set represents the current reality in the industry with many years of real-world inspection information; (2) it is available in large quantity as required by deep learning methods. The fact that this data set was not prepared for this experiment in mind presents some challenges but at the same time, brings with it the variability of assets, their failure modes, and the challenges of aerial viewpoints.

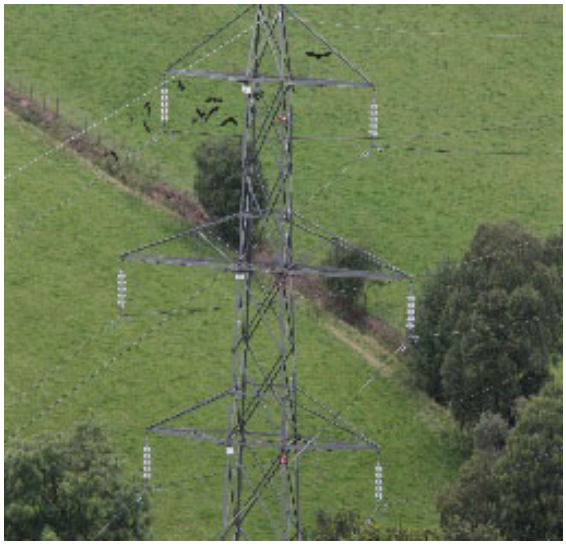
(a) Insulator Footing

(b) Insulator Top



(c) Insulator Middle

(d) Earth wire



(e) Top

(f) Middle

Fig. 3.1 Image categories selected for analysis.

(a) Footing (b) Anti climb guard

Fig. 3.2 A continuation of selected image categories.

(a) Whole (b) Signage

Fig. 3.3 Images excluded from the dataset

3.4 Image Data Preparation

Overall, 386,401 images of dimension $5616 \times 3744 \times 3$ pixels in RGB colours and in JPEG file format were provided. The images had been labelled to reflect the regions of tower with

Table 3.5 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

the condition parameter. A summary of image descriptions and the frequency of inspection is shown in Table 3.5.

The images had been subsequently analysed by experts in the industry and their collective condition ratings were allocated to towers. Each condition parameter was assessed based on a guideline previously agreed with the DNO. Image names were provided as hyperlinks on a spreadsheet. To inspect a tower, the images on disk were viewed via the hyperlinks. Extra information was provided as free text and some images of high-risk towers were provided.

Multiple images were provided for each tower with a single tower label (condition rating) was assigned. We assigned each tower label to all its images to use the images for supervised learning. Transferring tower labels to images enabled us to utilise an existing dataset for the investigation of paintwork, insulators, and U-bolts failure modes. If a tower was rated to be at high-risk due to paintwork failure, it is assumed that all images taken from that tower would have the same condition.

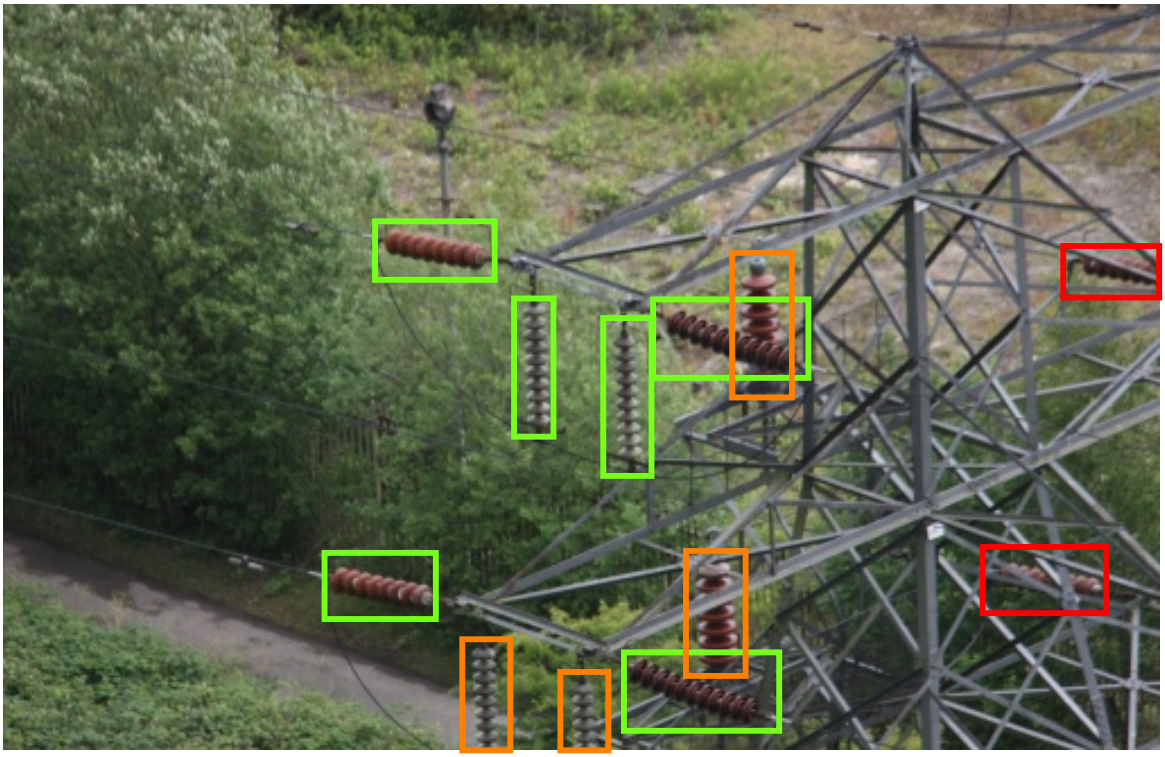


Fig. 3.4 Image annotation with different occlusion levels marked with red, orange and green boxes. Red: heavily occluded. Orange: partial occlusion. Green: clear components.

3.4.1 HDF5: Hierarchical data format

To train a deep learning network, data are fed in manageable batches. For small datasets, all data elements are moved to memory to speed up training. For image data, this is difficult as images take up memory quickly. So, a training pipeline usually allows fetching images from disk and that increases I/O issues. With images in several disk locations, training will suffer. To solve this problem, all the images and labels were serialised and saved in a hierarchical data format (HDF). This was done by first cropping the centre of the images to remove the meta data printed along the top of the photograph. Then, the images were resized to 256×256 pixels and added to the HDF database. Fetching data serially ensures that training can proceed with reduced I/O latency from disk/memory/GPU request cycle.

3.4.2 Annotation of images for object detection

Images were annotated to identify insulators and U-bolts sub-image regions using the VGG Image Annotator (VIA) [18]. Each object of interest was marked using rectangular bounding boxes and labelled as one of the following. Figure 3.4 illustrates the different levels of occlusion of insulators using colour coded boxes. Red means heavy occlusion, orange means partial occlusion and green means a clear view. The guideline used for detecting our objects of interest is shown below:

- Clear (green): 100% within the image with no occlusion.
- Partially occluded (orange): at least half of the object is within the image, but some part of it occluded either by the image boundary or by the tower or some other object.
- Heavily occluded (red): more than half of the object is occluded either by the image boundary or by the tower or some other object.

3.4.3 Preprocessing

Image preprocessing involves steps required to prepare image data for analysis. We used RGB mean subtraction and cropping of input images in some sections of this thesis.

RGB mean subtraction

Electrical overhead line networks are outdoor assets and during data collection, it would be difficult to control illumination. To normalise the varying lighting conditions, photographs were collected during spring and summer months. There is no guarantee that images acquired during spring and summer would have uniform brightness across the dataset. By subtracting the channel-wise mean of all the images from training data, the image appears uniformly darker and may be used to deal with the varying image brightness. In our experiment, we

Fig. 3.5 Random input crops of 224×224 pixels (dotted lines) from 256×256 pixels

subtract the mean of red channel, mean of green channel, and mean of blue channel from their corresponding channels in each image.

Cropping input image

As shown in Fig. 3.5, an image can be sampled by randomly cropping different $N \times N$ regions. The ability of a deep learning model to generalise well on targets it was not previously trained on depends on the quality and variability of the training data. There are several data augmentation methods for creating multiple copies of the same data point such that a model may learn more representative features and hopefully avoid over-fitting. The methods include image rotation, flipping and zooming. Cropping image patches is a useful technique for sampling input image and ensures that a network learns slightly different features of an image at every training step.

3.5 Dataset Limitations

The data set has the following challenges:

1. Images were taken from helicopters with human analysis in mind and human analysis introduced errors to tower labels.

2. The scale of objects in images are not uniform. Some images are close-shot, and others are long-range. Long-range photographs would affect fine-grained detection of components.

3. Objects in some images were occluded by tower structure or by other components.

4. Photographs of towers also contain roof-tops, vehicles, trees, etc. Extracting an object from a cluttered image is more difficult than on a clear view.

5. The occurrence of multiple targets in an image, e.g., a suspension and tension tower within the same field of view would affect detection of towers and consequently reduce assessment performance.

3.6 Data Subsets

All the images in T7799 were used in Chapter 4 without filtering image categories. The detection of specific sub-image regions would require images that are close range. Therefore, Chapter 5, utilised images from eight categories including Insulator Footing, Insulator Top, Insulator Middle, Middle, Footing, Top, Anti Climb Guard and Earth Wire. Images were selected from all three databases, i.e., T6775, T7638, and T7799 to ensure enough images that are likely to contain the objects of interest. These categories would also reflect the variability of components and views along OHL inspection corridor. Images of *whole tower* would be excluded because they were taken at long-range and detecting specific components would require extra zooming. Also, images of *signage, cable platform, substation, and spacer* (examples in Fig. 3.3) were not considered as they would not contain the objects of interest. The remaining eight categories constitute 94.2% of images and are those likely to

contain the objects of interest for detection of objects. Together they should be representative of the variability encountered along the OHL inspection corridors. Data sets were sampled from this set of 363,859 images for the detection of components and classification of towers.

For tower condition classification, condition labels were provided for left and right circuits as separate assets but objects in the image were not labelled to identify the circuits. As a result, the labels provided for left and right circuits in the data set were aggregated by a max operation to produce a single label.

Chapter 4

Classification of Tower Configurations

4.1 Introduction

Overhead lines transmission and distribution networks are typically deployed as straight lines e.g., from point A to E as shown in Fig. 4.1. However, due to some barriers on the line of travel, the ideal straight powerline is difficult to achieve. Using Fig. 4.1 as an illustration, the network deviates by an angle t at C and could then navigate through D to E before they can continue on the straight line. The amount of deviation on a line segment determines the type of tower installed at each point on the network. That also means the types of components used on each tower type vary. If the angle of deviation, $t < 2°$, a suspension (S-type) tower is used and tension (T-type) when the deviation, $t > 2°$ [91]. At the point of deviation, towers structures are properly re-enforced to withstand the tension force exerted on the cross-arm by the conductors. T-type towers have more complex designs and pulls on cables as they negotiate a bend on the segment. S-types are simpler tower structures that hold cables in a suspended position. In general, tower configuration is dependent on how cables are deployed on them.

Some examples of suspension (S-type) towers are shown in Fig. 4.2. These are characterised by both tower structures and how cables are suspended on it. Insulators on S-types

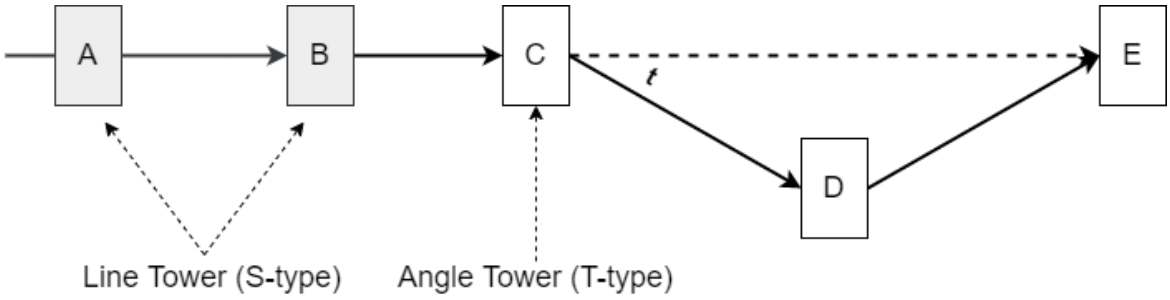
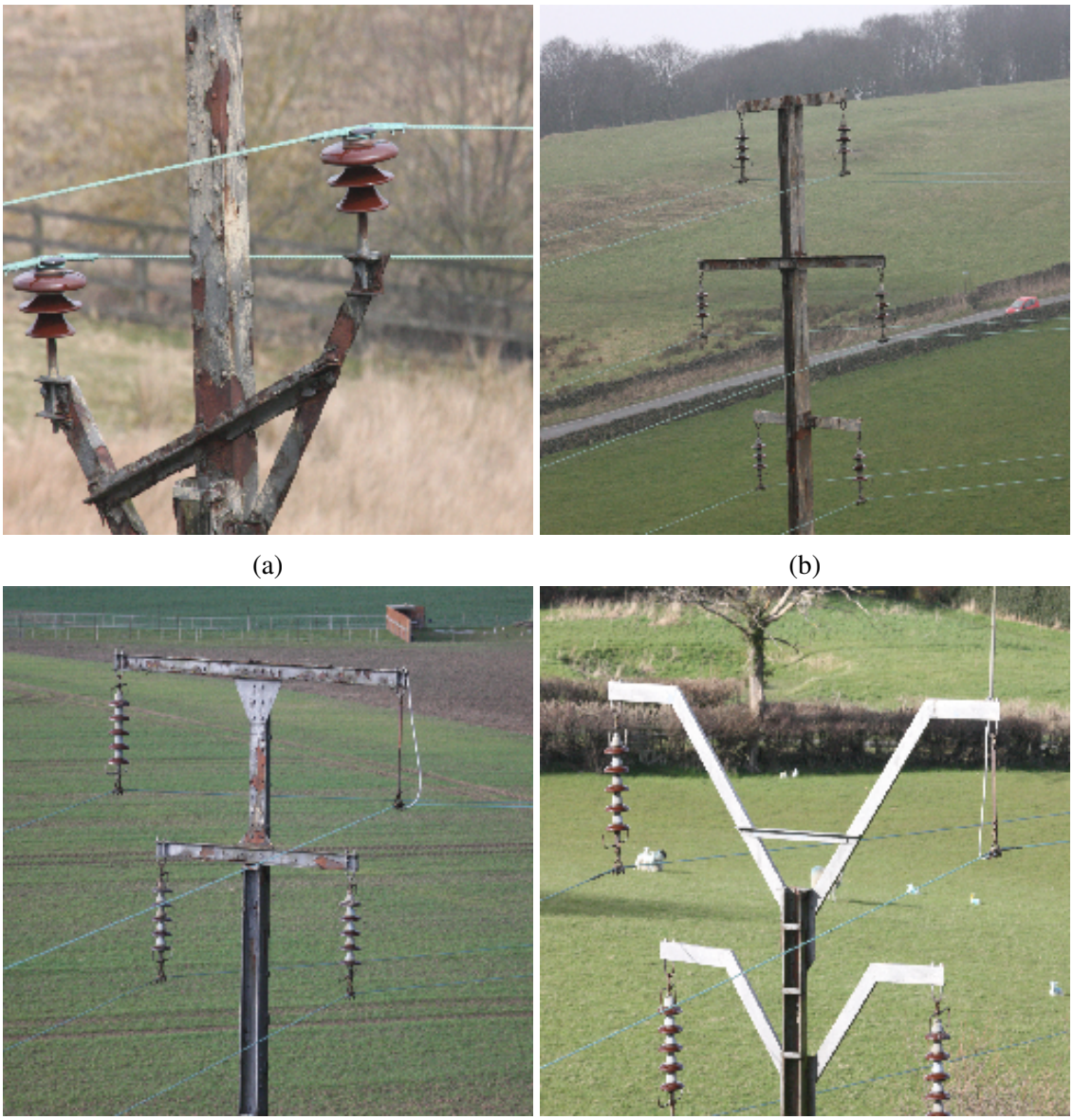


Fig. 4.1 Schematic diagram of a line segment showing positions of line (suspension) and angle (tension) towers.

are usually perpendicular to the line (cable) either below or above. In the design in Fig. 4.2a, cables are supported above the insulator. Whereas in Fig. 4.2b, 4.2c and 4.2d, cables are below the insulators. So, the structural resemblance does not determine the tower configuration. These are line towers used typically at points A and B of Fig. 4.1. Fig. 4.3 shows some examples of tension (T-type) tower structures typically used at points C, D and E of Fig. 4.1. A typical T-type tower has more re-enforcement and can bear the load at those points on the line.

As stated previously, where a tower is used on the network determines the component they support. That also means, inspection parameters on a tower are often associated with the tower type. For example, insulators and fittings vary between configurations. To be able to assess components of the tower, there is need to first detect and classify tower types. The classification of tower configuration serves as a precursor for detecting specific components during an inspection process.

This Chapter describes a novel application of deep learning for tower classification. Specifically, towers are classified as being either S-type or T-type. Tower classification was achieved by first classifying each of the multiple images of a tower and then using voting to determine the tower class. Since many images do not contain the relevant parts of the tower, a third image class was introduced, i.e., the unknown (U) image class. We trained 3-class classifiers to predict S, T or U image types. Images of the unknown configuration



(a) (b)

(c) (d)

Fig. 4.2 S images with different structural designs

are shown in Fig. 4.4. In subsequent sections of this thesis, tower components around the cross-arm will be investigated and classifying towers as S-type and T-type will be useful in such applications. Tower components localised around the U region such as concrete muffs, safety signs, and their conditions will not be covered. In addition to the classification of towers, this Chapter demonstrates that the identification of U images is feasible.

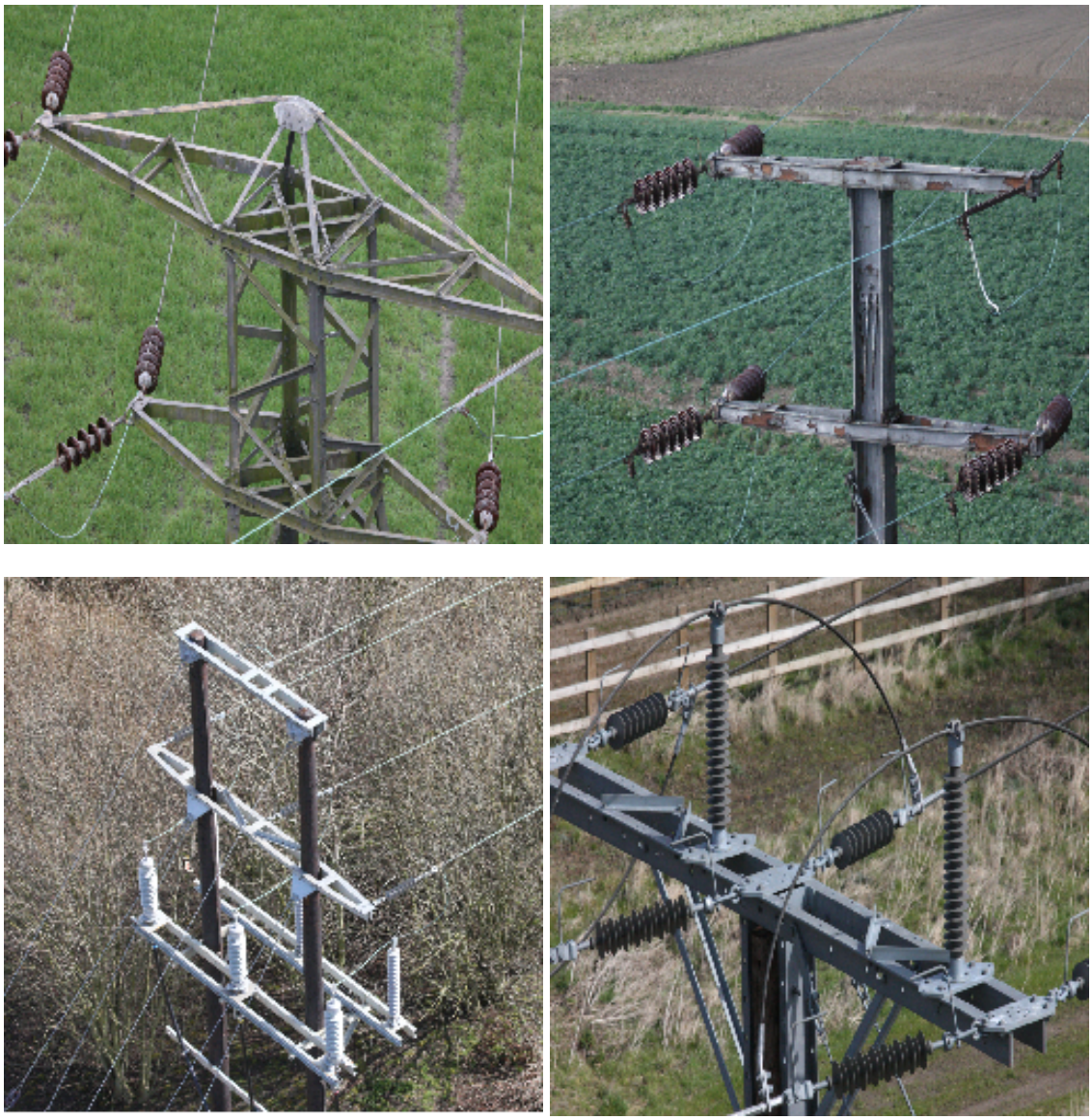
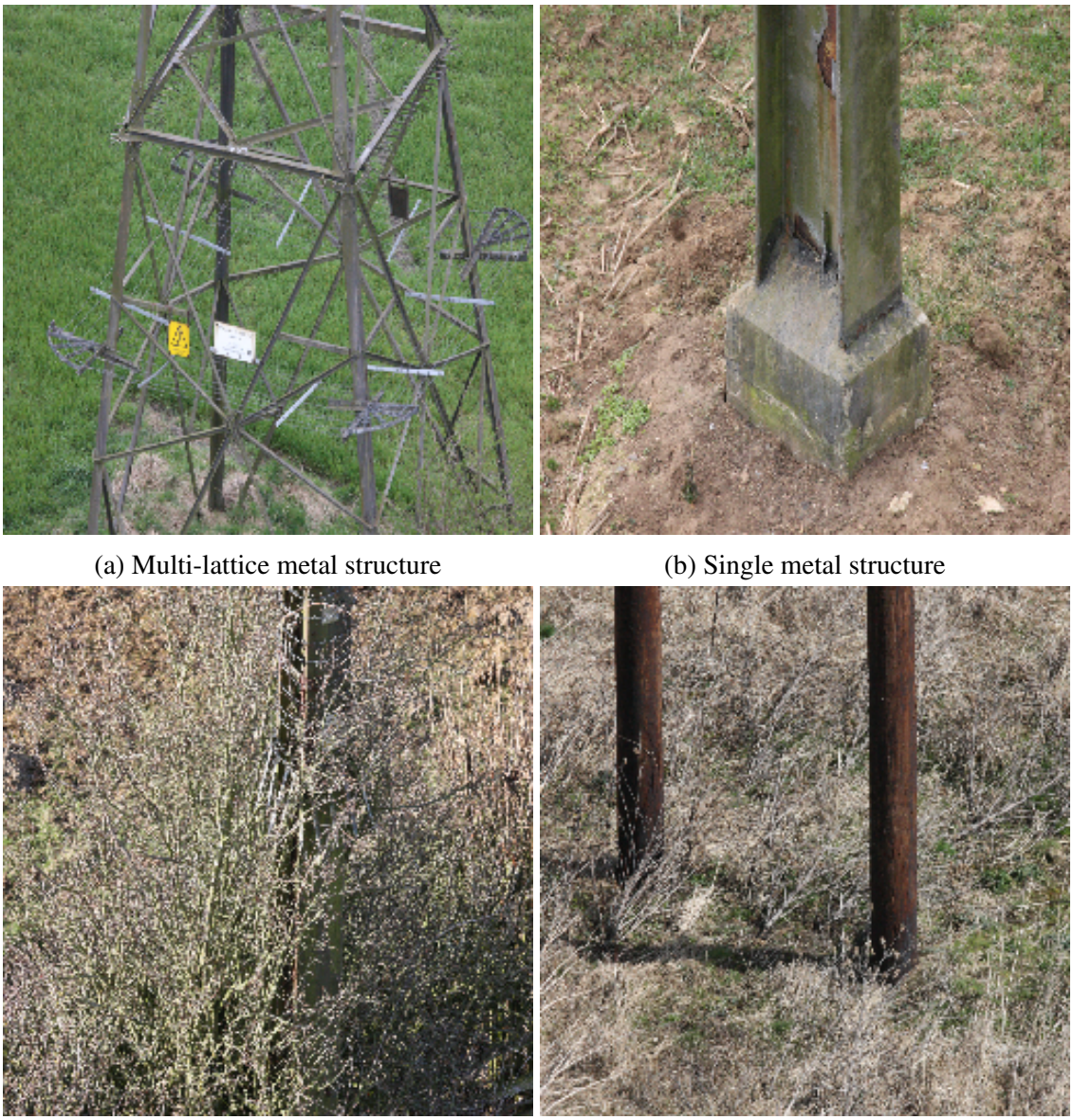


Fig. 4.3 T images with different structural designs

4.2 Experimental Methods

A typical image classification problem assigns one label per image. Tower images used for this experiment were selected from T7799 (Appendix A) comprising of medium and low voltage lines. Towers were presented as multiple images with a label assigned to each tower bag. Different categories of images contained in each tower bag was described in Section 3.4. Images of cross-arm regions such as those shown in Fig. 4.2 for S-type and



(a) Multi-lattice metal structure

(b) Single metal structure

(c) Single metal structure

(d) Wooden poles

Fig. 4.4 Images in which tower type is not apparent (U)

Fig. 4.3 for T-type towers would be informative for classifying tower configurations but those of the body/footing regions of a tower (Fig. 4.4) will not. Images of tower body or leg regions do not have cross-arms or the components around the cross-arm such as insulators and other fittings. Images of tower body and footing regions were labelled as U (unknown) configuration because their tower type is not apparent. At image level, there are thus three

Table 4.1 Distribution of data for training, validation and testing sets.

	Tower label	Number of towers	Image label	Number of images
Training	S-type	519	S	12,621
			U	4,282
	T-type	270	T	9,166
			U	2,163
Validation	S-type	79	S	1,963
			U	739
	T-type	41	T	1,469
			U	422
Testing	S-type	80	S	1,829
			U	837
	T-type	38	T	1,220
			U	359
Total		1,027		37,070

class labels: S, T, and U. To avoid bias, all images from a tower must be used for either training or testing. Therefore, data splitting was at tower level. There are 28,232 and 4,593 images for training and validation, respectively.

ResNet50V2 [36] and VGG-16 [112] networks were tested for the classification of OHL tower images as S, T, and U. The selected deep learning architectures are standard convolutional neural networks (CNN) and are a good choice for extracting features from images. The models are available in Keras. Keras is an application programming interface (API) for deep learning. There are pre-trained networks within this API, which can be used off-the-shelf. Networks pretrained on ImageNet dataset are typically used because of the large size and contains everyday objects. The pretrained models are easy to use as starting points for learning new representations with little modification. In this experiment, transfer learning approach was used for training the network. Transfer learning is a technique in deep learning where the features learned in one application are used for another application. This was used here because of the relatively small number of tower images selected for this experiment. Also, our data set, like all aerial images is characterised by background clutter

comprising of roof-tops, cars, humans, etc. Some of these background objects in our data set may have been previously learned from ImageNet data set. By re-training such a classifier on a data set of tower images with this kind of background noise, the network might perform better by building on existing weight (the known). In addition to faster training, this can boost performance and generalisation [126].

As shown in Fig. 4.5a, the selected networks were modified to have three outputs (3-classes) instead of 1,000 classes of ImageNet and fine-tuned on tower images. Specifically, the penultimate layer was flattened and connected to a new 512 nodes, which then feeds a 3-nodes FC layer and a Softmax activation. The 3-nodes correspond to the number of classes on the current task. There is an approach to transfer learning known as fine-tuning. This involves re-training a network to learn new weights while preserving previously learned weights. The base network i.e., one-half of the network was kept constant for 15 epochs, while the other half trained on the new data to learn new features without destroying previous knowledge. Then, the whole network was trained further for up to 50 epochs. Table 4.2 shows a summary of the selected training parameters.

Training used multiple images of each tower, followed by voting to determine tower configuration. Each input image has 244×244 pixels as required by the networks. Standard image augmentations were applied using Keras image data augmentation function. The augmentation was done online during training, and it creates modified versions of input images by randomly zooming, vertical shifting and horizontal flipping of inputs. It ensures that the models do not train on the same set of images at every step. Fig. 4.6 shows examples of the effect of horizontal flipping on original image. Image augmentation is useful to avoid over-fitting. Stochastic Gradient Descent (SGD) with learning rate of 0.001 was used. The categorical cross entropy loss function was used for training. A batch size of 32 images was used and found to have worked well for this data set. However, [54] concluded that a higher

Algorithm 1: Image-based classification of tower configuration

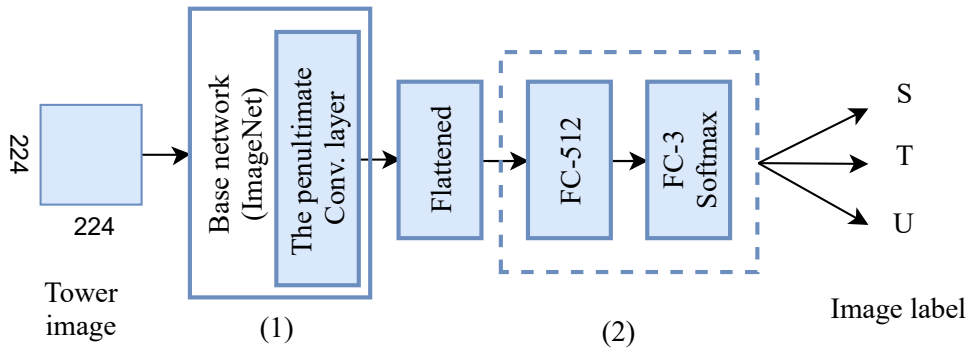
```
   Data: tower
   Result: config
1  S ← 0; T ← 0 ;                                   /* Initialisation */
2  while img ∈ tower do
3      image ← imageRead(img) ;   /* Read an image of 224 × 224 pixels */
4      y ← f(image) ;             /* A classifier with Softmax output */
5      label ← argmax(y) ;   /* Index with the highest class probability
        (i.e., S, T or U) */
6      if label := S then
7          S ← S + 1;
8      else if label := T then
9          T ← T + 1;
10     else
11         Don't care
   ;       /* Return the label with the higher number of occurrence */
12 if S > T then
13     config ← S ;                          /* config is an S-type tower */
14 else if T > S then
15     config ← T ;                           /* config is a T-type tower */
16 else
17     Refer to operator when there is a Tie
```

batch size does not usually achieve high accuracy. Steps for evaluating the method is shown in Algorithm 1.

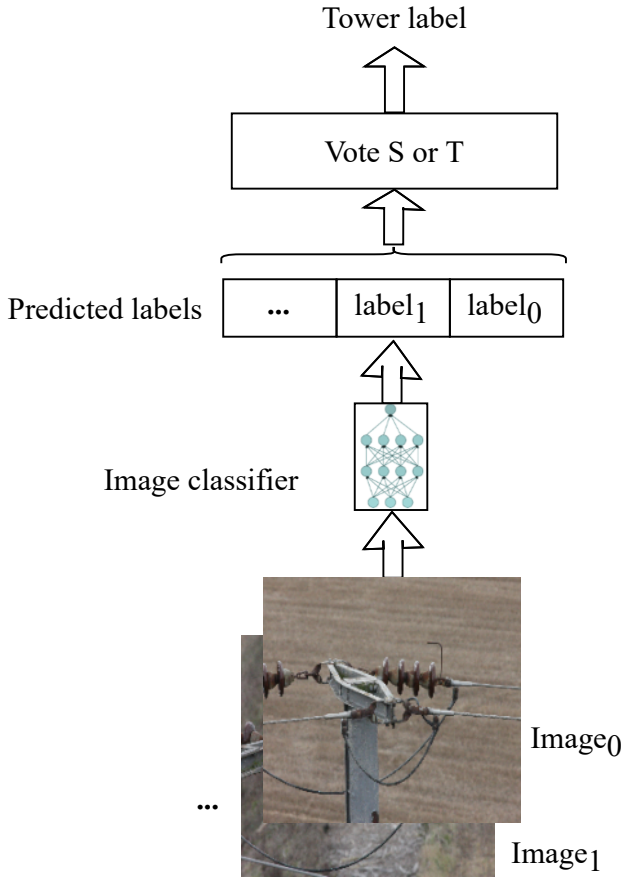
4.3 Results

4.3.1 Image level classification

There are 118 towers and 4,245 images of S, T, and U in the test set. The distribution of towers and images for testing is shown in Table 4.1. The VGG-based classifier predicted 97.04%, 97.69% and 96.32% of S, T and U test images correctly. The ResNet classified 96.99%, 96.54% and 95.65% of S, T and U test images correctly.



(a) Image level: (1) The base network using either a ResNet50V2 or VGG-16 previously trained on ImageNet. Output of the penultimate Conv. layer was flattened and connected to (2) newly added set of fully-connected layers, Softmax output with 3-nodes for S, T, and U.



(b) Tower level: Classify all the images of a tower followed by voting to determine tower type.

Fig. 4.5 Classification of towers by function

Fig. 4.6 Image augmentation. Top: Original images. Bottom: Horizontally flipped images

Table 4.2 Network parameters. During the first 15 epochs, back propagation is limited to the new fully connected (FC) layers.

Parameter	Value
Size of input image	224 x 224
Number of images per batch	32
Maximum number of epochs	50
Optimisation function	Stochastic Gradient Descent (SGD)
Learning rate	0.001
Loss function	Categorical cross entropy

Fig. 4.7, 4.8 and 4.9 show some examples of S, T, and U images, respectively, that were correctly classified. Comparing the correctly classified images with those that were missed, one notices that close-range images with relatively clean backgrounds were correctly classified. Fig. 4.10, 4.11 and 4.12 show examples of wrongly classified S, T and U images. These figures highlight some characteristics that affect correct detection including:

- Long-range images.
- Heavy background clutter, e.g., houses, trees.
- Instances of multiple objects.
- Cases of wrong labelling.

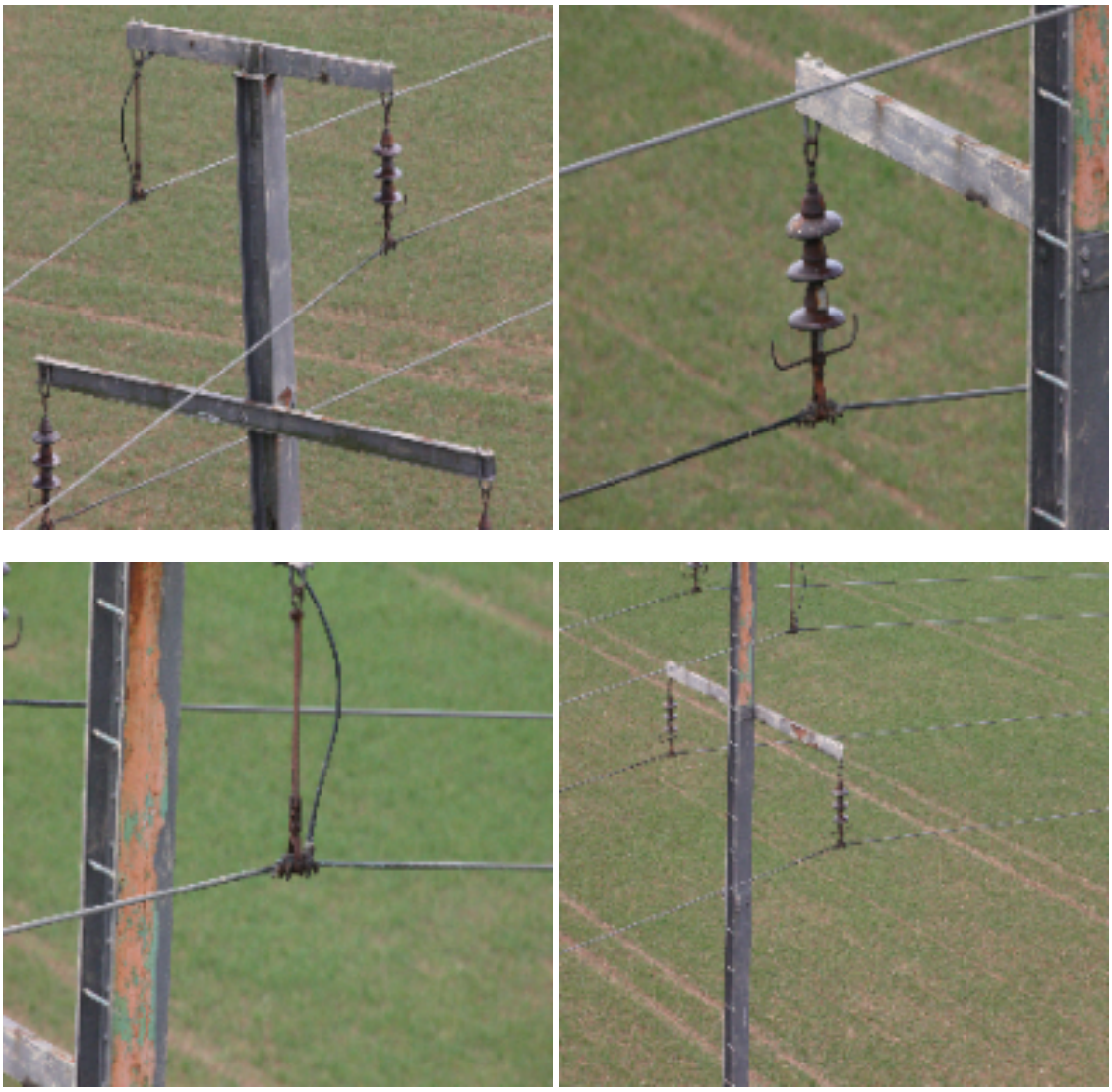


Fig. 4.7 S images correctly classified

4.3.2 Tower level classification

An aim of this study is to classify towers as suspension (S-type) or tension (T-type). Each tower was provided as a bag of images. Within each tower bag in this data set, there are 32 images on the average. Algorithm 1 and Fig. 4.5b show that simple majority voting on multiple images of each tower was used to determine tower label. This samples all the image predictions for each bag and counts the number of occurrences of S and T labels. The label with the highest count is returned as the final prediction for the bag (i.e., tower

Fig. 4.8 T images correctly classified

level classification). Images classified as U are discarded because tower configuration is not apparent in them. A situation may arise in which there is a tie (equal predictions of targets). However, as shown in Table 4.3 as well as Table 4.4, there was no tie in our experiments. In the event of a tie, a secondary examination by human inspection could be used to correct the error or by referring such a case to asset owners for more information. To avoid occasions of a tie, an alternative method should be used. Popular methods of aggregating multiple image predictions include averaging and using the maximum of the set of predictions.

Fig. 4.9 U images correctly classified

As shown in Table 4.3, the VGG-based classifier classified 117 towers correctly and misclassified one S-type tower as T-type. On the other hand, ResNet-based classifier predicted

Table 4.3 Confusion matrix of tower level classification: VGG16-based

		Predictions	
		S-type	T-type
Actual	S-type	79	1
	T-type	0	38

(a) System limitation

(b) Heavy background clutter

(c) Multiple objects in the image

(d) Multiple objects in the image

Fig. 4.10 S images misclassified as T

Table 4.4 Confusion matrix of tower level classification: ResNet-based

		Predictions	
		S-type	T-type
Actual	S-type	80	0
	T-type	0	38

all 118 towers correctly. The confusion matrix of ResNet-based classifier tower classification is shown in Table 4.4.

(a) System limitation (b) Long-range image

(c) Clutter and Long-range image (d) Long-range and multiple objects

Fig. 4.11 T images misclassified as S

4.4 Summary

Manual classification of towers from images can be challenging. As shown in Fig. 4.12, some images were wrongly labelled. This a good example of human limitations. A human classifier can become tired of the repetitive task, and this can result to errors. The methods presented in this Chapter are robust for the classification of tower configuration. Fig. 4.13 and Fig. 4.14 show the Grad-CAM visualisation of tower type predictions. The heat maps suggest

(a) Heavy background clutter

(b) Wrong label

(c) Wrong label

(d) Wrong label

Fig. 4.12 U images misclassified as T

the parts of a tower necessary for predicting both T-type and S-type tower configurations. Insulator regions appear to be informative for the task.

Correct classification of tower images is subject to several factors as highlighted in Fig. 4.7, 4.8 and 4.9. Some of the characteristics that may have helped in correct predictions include (1) single object within an image, (2) clear background and (3) close-up images. These are examples of how images should be presented for automated visual inspection of

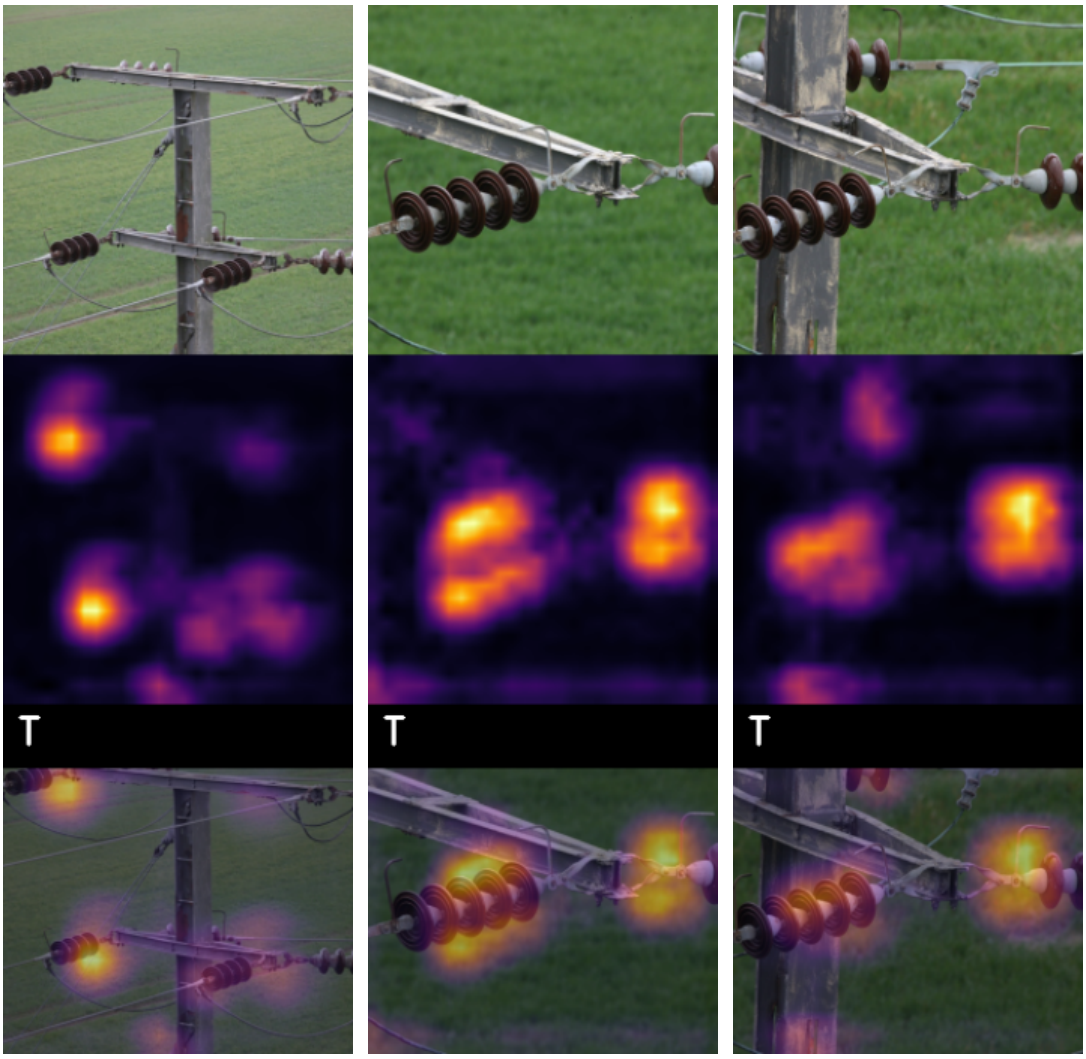


Fig. 4.13 Grad-CAM visualisation of T tower images. TOP: RGB images (input), MIDDLE: corresponding heat maps, BOTTOM: Heat maps on the input image to visualise regions of interest.

OHL assets from aerial images. These features are not always guaranteed as reflected in the misclassified images showing (1) long-shot (2) images with multiple objects within the field of view, etc. Images used in this experiment were taken from helicopter, which explains some of the limitations observed in the data set. There are alternative aerial platforms such as unmanned aerial vehicle (UAV), which could be used for sites that are difficult for the helicopter.

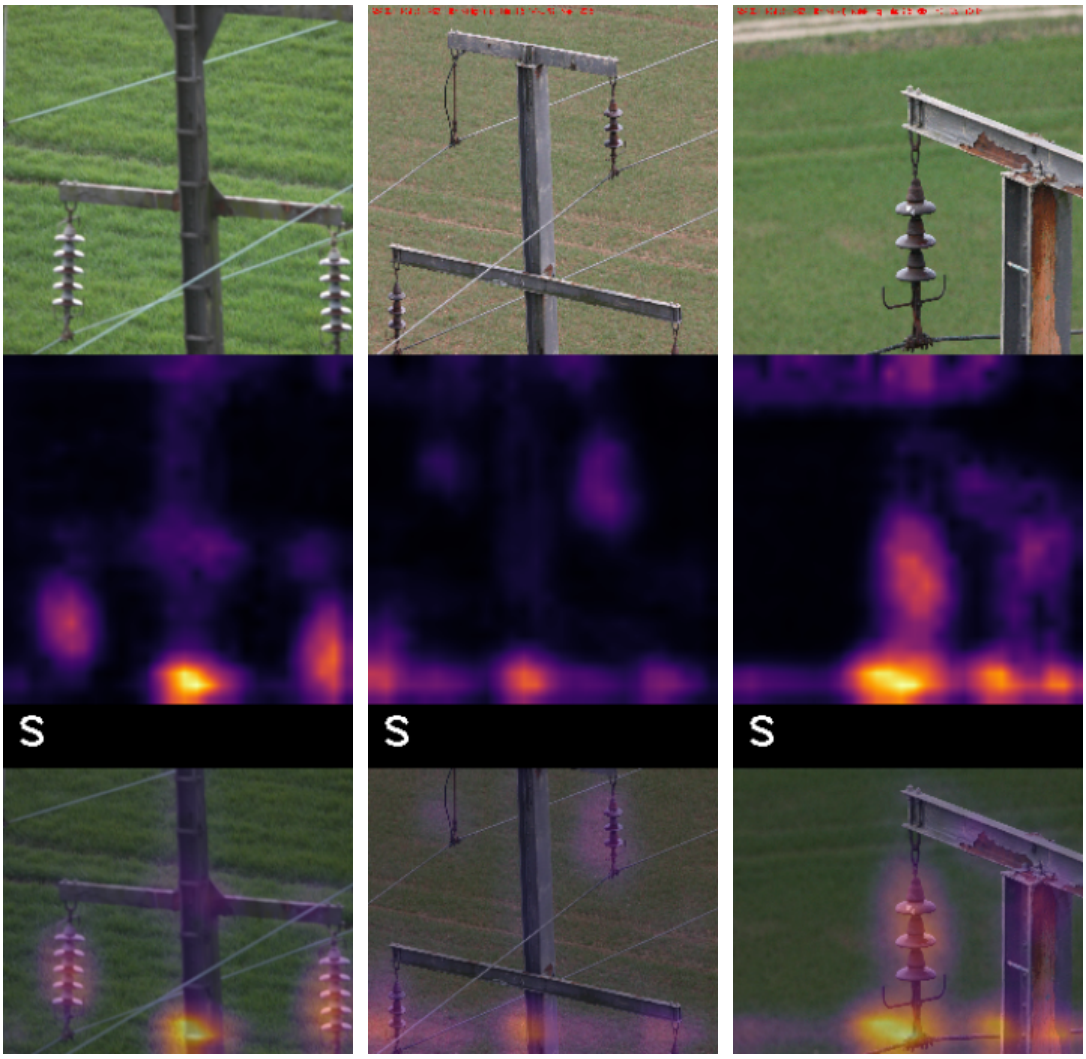


Fig. 4.14 Grad-CAM visualisation of S tower images. TOP: RGB images (input), MIDDLE: corresponding heat maps, BOTTOM: Heat maps on the input image to visualise regions of interest.

This work represents the first report of deep learning-based classification of tower as suspension or tension. In other words, towers were classified by function on a network segment. In contrast, [102] presented the classification of towers using the resemblance of pylon structures. The industry requires the classification of towers according to how they are used on network segments. The function of a tower determines its configuration and the components they support. This Chapter demonstrated the effectiveness of deep learning-based methods for the task. Successful classification of tower configurations would serve

as a precursor for the detection and assessment of other overhead line assets. In subsequent Chapters, components like insulators and tower attachments such as U-bolts will be detected for condition assessment. Whereas insulators can be detected in both S-type and T-type towers, insulator U-bolts are visible on S-type alone. The detection of assets such as concrete muffs, DODs, tower name plates, etc. would be possible with the detection and classification of tower body and leg regions. Components around tower body and leg regions will not be covered in this thesis. However, the assessment pipeline could be extended to assets in that region as well.

Chapter 5

Detection of Insulators and U-bolts on Electricity Tower Images

5.1 Introduction

Power outages along overhead line (OHL) networks are caused by several factors including weather, vegetation encroachment, activities of birds, faulty equipment, or human error. Haes Alhelou et al. [32] reported that 31.8% of the blackouts world-wide between 2011 and 2019 were caused by faulty equipment and/or human error. The monitoring of faults along the power line corridor presents a huge financial commitment to network operators. In the UK, distribution network operators (DNOs) are required by law to inspect OHL assets with sufficient frequency and to maintain a record of such inspection for a period of not less than 10 years [120]. The regulation aims at preventing the events of blackouts and the unprecedented, cascaded effects including economic losses. OHL assets are typically inspected every 2 - 3 years. The frequency of inspection amounts to high recurrent expenditure considering the span and quantity of components on the networks.

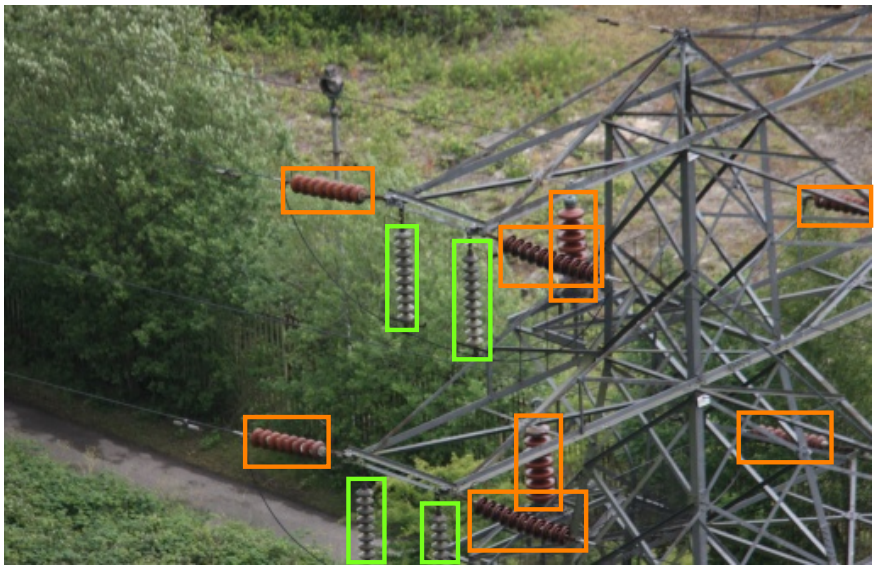
The state-of-the-art for inspecting electricity network assets uses images from manned aerial vehicles. Several images are taken and analysed off-line by experts. Experts would

search each tower image for insulators and other condition parameters like U-bolts and vegetation cover. There are over 50 condition parameters for measuring the status of a line on the network. The current inspection pipeline is labour intensive, costly, and subjective. Results of assessment by a human expert depend on training and level of experience in the field. As stated in section 4.4, the function of a tower determines its configuration and that determines the components deployed on them. Even among the same type of tower, the components extracted are so varied based on colour, orientation, etc and these impact inspections. As a result of these challenges, efforts to automatically extract these components of interest are on the increase to facilitate condition assessment and management.

There is a wide spectrum of assets on OHL networks to be detected. The major assets of interest usually are towers, insulators, conductors, and U-bolts. Among these components, insulators are the most explored [65]. Insulators are more frequently inspected because of the critical role they play in ensuring the safety and integrity of the network. Insulators prevent the leakage of power. In addition to financial losses, power leakage could be dangerous to members of the public and any life coming in contact with the towers. The detection of insulators from aerial images is a challenging task for several reasons. The most challenging is that electricity networks are in open environments and thus prone to varied viewing conditions and clutter [33, 81, 84]. Figure 5.1a shows a typical tower image with very complex background. The object in view in this image is the tower but at the background are roof-tops, vehicle and most importantly, other lookalike assets.

Poor contrast affects detection of components e.g., detecting glass insulators. In addition, there are several designs of insulators such as ceramics, glass, porcelain, and composite types. They also come in different dimension, shape and colour. Some authors appear to limit detection of insulators to subsets, e.g., glass only, or glass and porcelain [24, 56, 61]. However, it is common for a high voltage tower to be supported using different insulator strings. Fig. 5.1b shows a typical tower supported by two types of insulators (glass and

(a) Typical example of background clutter

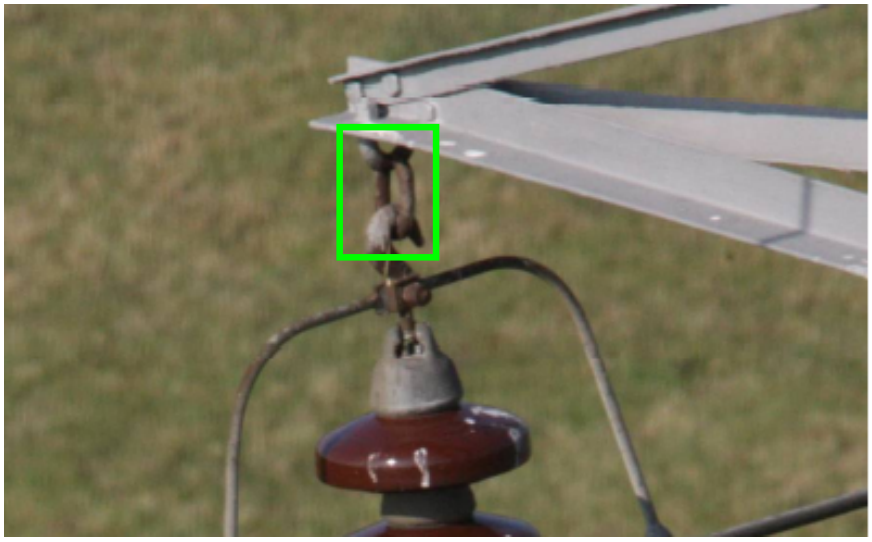


(b) Typical example of a tower supported by mixed insulator types. Green denotes glass and orange denotes porcelain type.

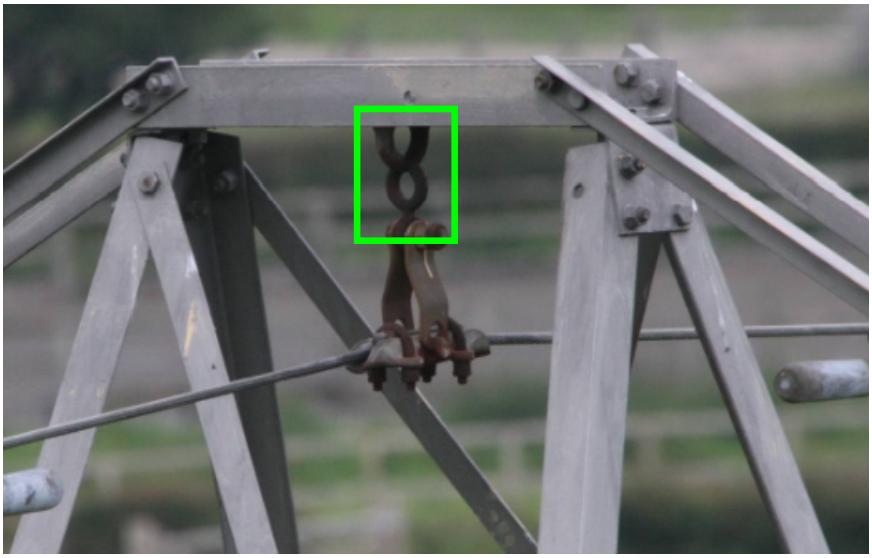
Fig. 5.1 Illustration of (a) complex background (b) variability of objects on tower.

porcelain). These multiple insulator types with varying size, orientation and colour need to be detected simultaneously.

Another asset identified as a high-risk condition on a tower is the U-bolt. The U-bolt is a 'U' shaped structure usually attached to the tower, and it supports insulators and earth wires on suspension towers. Fig. 5.2 highlights regions of insulator U-bolt and earth wire U-bolt.



(a) Insulator U-bolt shackle



(b) Earth wire U-bolt shackle

Fig. 5.2 Types of U-bolts

Compared to insulators, U-bolts are smaller and do not have the distinct colour feature like insulators. Generally, weather, and other environmental factors affect electrical components leading to different kinds of failures. Moisture accelerates corrosion and erosion of U-bolt parts. Hence, early detection of this failure mode is very important.

The focus of this Chapter is to localise the objects of interest, which is a first step for assessing their conditions. In addition to insulators, this Chapter demonstrates the detection of U-bolts from aerial images. The main contributions are:

1. Towers are sometimes supported by different insulators and our real-world inspection data set is representative of the diverse component types, scenes, and views. We demonstrated that multiple insulator types i.e., glass, porcelain, ceramic and composite can be detected simultaneously.

2. Detection of U-bolts from aerial images. This represents the first study on automated detection of U-bolts from aerial images.

3. Comparative evaluation of Mask R-CNN and RetinaNet for electricity network components identification.

5.2 Experimental Methods

This section of the thesis describes the use of two state-of-the-art networks for insulators and U-bolts detection. Mask R-CNN [35] and RetinaNet [59] were fine-tuned on aerial images of electrical overhead line towers. In the series of region-based detectors, Mask R-CNN is more recent. RetinaNet was designed to bridge the gap between one-stage and two-stage networks by eliminating the bottleneck posed by background-foreground class imbalance [59]. These are state-of-the-art detectors and are available in Keras for use as starting point for object detection. The implementations used were based on Resnet backbone and were previously trained on the COCO data set [60]. Also, the networks have not been previously comparatively analysed for this application. Fig. 5.3 shows a block diagram of instance detection subsystem.

Considering the size of the data sample used, we fine-tuned the networks on tower images to learn new representations. Fine-tuning means re-training a network on a new task while

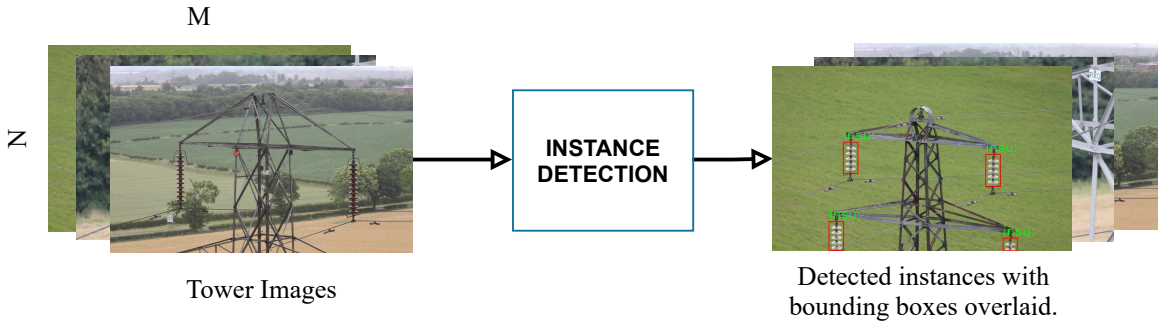


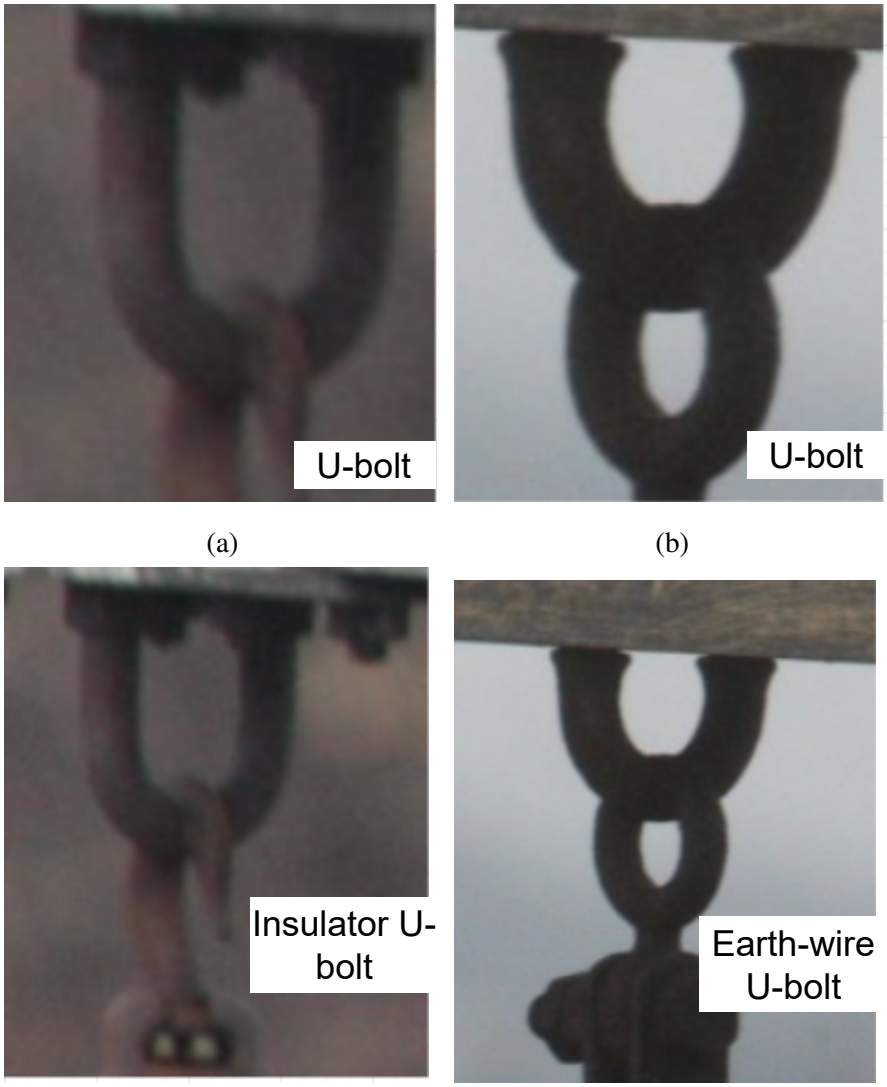
Fig. 5.3 Instance detection: Mask R-CNN and RetinaNet, which have been previously trained on COCO images detect insulators and U-bolts from tower images of $M \times N$ pixels.

Table 5.1 Parameter selection and the values used for training the detection networks.

Parameter	Mask R-CNN	RetinaNet
Initial learning rate	0.001	☑
Backbone Network	ResNet101	ResNet50
Size of input image	1024 x 682	☑
Images per GPU	1	2
Maximum number of epochs	35	50
IOU threshold	0.5	☑

preserving previously learned information (weights). Training and testing of these networks were carried out on a RTX2080 GPU with 10GB memory. Table 5.1 shows basic training parameters selected for the networks and values used. Whereas Mask R-CNN was trained using a batch size of one image of size 1024×682, RetinaNet was train with two images of the same size per batch. The output layer was trained for the first 10 epochs using a learning rate of 0.001. Then, the entire network was trained for between 35 and 50 epochs. While training the entire network, the learning rate was changed to 0.0001.

Furthermore, RetinaNets were trained on single classes and compared to another trained as a multi-task detector. As shown in Fig. 5.4 (c and d), U-bolts can be categorised into insulator U-bolts and earth wire U-bolts. DNOs inspect these components separately in terms of condition rating. However, insulator U-bolts and earth wire U-bolts are identical in structure Fig. 5.4 (a and b). While insulator U-bolt support insulators, earth wire U-bolt's



(a) (b)

(c) (d)

Fig. 5.4 Automatically extracted U-bolts sub-image regions

support earth wire clamp. To automate the detection of both insulator U-bolts and earth-wire U-bolts, more pixels were added around the region of interest to introduce some contextual information. Hence, the classes for our detector were insulator, insulator-U-bolt, and earth wire U-bolt.

The detection models were evaluated using average precision (AP) and compared using their response time. AP is a measure of the performance of a detector determined by computing the area under the precision and recall curve. First, an intersection over union

(IOU) was computed as a measure of the overlap between the ground-truth bounding box and the detected bounding box. As shown in Fig. 5.5, the level of overlap varies among the objects in the image resulting to values between 0 and 1 with 1 being full overlap. In this experiment, IOU was set to 0.5. This means that for IOU ≥ 0.5, a detection would be considered a true positive (TP) detection. True positive rate is the proportion of positives that were correctly detected out of the total positive examples. Other parameters for evaluating a detection system include the false-positive (FP) or false-alarm. FP is the detection of objects that have zero overlap with ground-truths, e.g., detected background regions. Missed ground-truths are considered false-negatives, FN. Precision, P is defined as the ability of a model to predict only the relevant objects and computed as the percentage of correct positive detections over all detections. Recall, R, determines the percentage of true positives to all ground truths. P and R were calculated as shown in Eqn. 5.1 and Eqn. 5.2, respectively. We evaluated the models using average precision (AP) [87].

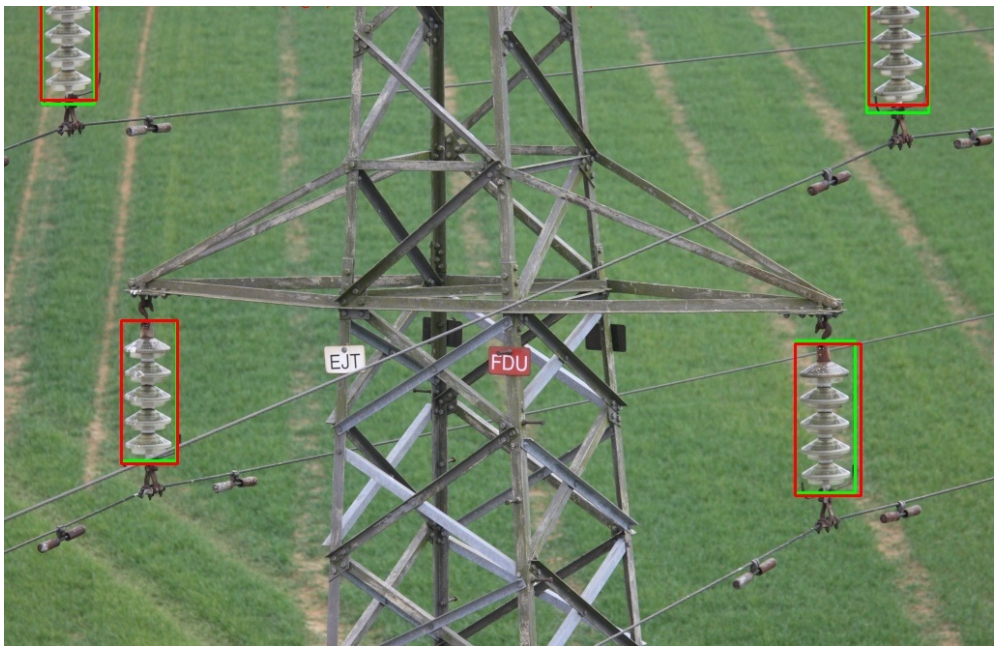


Fig. 5.5 Overlap of ground-truth and detected bounding boxes. Green denotes ground-truth bounding boxes. Red denotes detected bounding boxes

$$P = \frac{TP}{TP + FP} = \frac{TP}{\text{All detections}} \tag{5.1}$$

$$R = \frac{TP}{TP + FN} = \frac{TP}{\text{All ground-truths}} \tag{5.2}$$

5.3 Results

Mask R-CNN and RetinaNet have been used for different applications in this domain but have not been compared as presented in this study. For example, Mask R-CNN was used in [48] for insulator detection and RetinaNet in [121] for pin detection. The detection networks were tested as follows:

1. Detection of four insulator types simultaneously. Insulators are varied in scale, orientation and colour.
2. Determine the effect of occlusion on detection.
3. Detection of U-bolts on aerial images.
4. Detection of sub-classes of U-bolts i.e., Earth-wire U-bolts and insulator U-bolts.
5. Compare single and multi-task object detection on objects of electrical tower.
6. Determine how object scale affect the detection of insulators and U-bolts; U-bolts are small objects compared to insulators.

5.3.1 Detection of insulators in randomly selected tower images

Eight image labels were selected for this task including images of *Insulator Footing, Insulator Top, Insulator Middle, Middle, Footing, Top, Anti Climb Guard and Earth Wire*. The selected categories were labelled to reflect the region of tower in view and the number of images is

Table 5.2 Insulator detection in randomly selected towers using Mask R-CNN

Occlusion Levels (see section 3.4.2)	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

representative of the frequency of inspection of the tower regions (Chapter 3). A total of 1,600 images comprising of 200 images from each of the eight image categories were picked at random. This was split into 75% and 25% for training and testing, respectively. Insulator regions were labelled on 1,200 images used for training and 400 used for testing. A total of 3,236 insulators were manually identified in 833 images of the training set, while 367 of the selected images of the training set did not contain insulators. There are 949 insulator regions in the test set of which 115 were heavily occluded.

As shown in Table 5.2, Mask R-CNN was tested with different image sizes. Inputs of 1024×682 obtained 83.8% AP on the overall test examples and improved to 89.4% AP with heavily occluded instances removed. When tested using images of dimension 512×341 i.e., reducing input size by half, performance dropped by approximately 3%. Objects become less resolved farther away from the point of view and were missed by the detector.

5.3.2 Detection of insulators and U-bolts in high-risk towers

Inspection reports for each route are usually accompanied by additional free text comments of the major concerns. Towers reported in the comments have high-risk conditions and require immediate attention including repairs or replacements of affected components within the inspection window. Some image names were provided in the free-text (up to four for each tower), which were used in this experiment. These images were used for the following reasons:

Table 5.3 Insulators and U-bolts from high-risk towers

	Towers	Images	Insulator	U-bolt
Training	598	1,466	1,966	1,580
Testing	150	364	496	382
Overall	746	1,830	2,462	1,962

- Helicopter-based inspection of high voltage lines take photographs at reasonable distance from the tower mainly for safety reasons. The resulting data set is characterised by many long-shot images. Objects are less resolved farther away from the camera, and this impacts detection.

- To understand the characteristics of the images upon which human experts based their judgement and to learn how these would improve the performance of automated detection.

A total of 1,830 images were extracted from 746 towers. As shown in Table 5.3, 1,466 images (598 towers) were used for training. From these images, 1,966 insulators and 1,580 U-bolts sub-image regions were manually annotated. The test set comprises of 364 images from 150 towers in which 496 insulators and 382 U-bolts regions were manually identified.

As shown in Table 5.4, with input size of 1024×682 pixels, Mask R-CNN detected insulators at 96.7% AP and U-bolts at 96.0% AP. Using a smaller input size of 512×341 pixels, the network recorded 95.6% AP and 93.1% AP for insulators and U-bolts, respectively. For inputs of 1024×682 pixels, RetinaNet detected insulators at 96.4% AP and U-bolts at 86.9% AP. The average inference times were $157ms$ and $60ms$ for Mask R-CNN and RetinaNet, respectively. Fig. 5.6 shows the Precision-Recall curves for Mask R-CNN and RetinaNet for insulator detection. With properly resolved objects in the image, the result of RetinaNet is comparable with Mask R-CNN while 3x faster. For smaller objects like detecting U-bolts, Mask R-CNN shows better capability (Fig. 5.7).

Table 5.4 Detection results in high-risk towers

Model	Time (ms)	Object	Input size	AP_{50}
Mask R-CNN	157	Insulator	1024×682	96.7
		U-bolt	1024×682	96.0
"		Insulator	512×341	95.6
		U-bolt	512×341	93.1
RetinaNet	60	Insulator	1024×682	96.4
		U-bolt	1024×682	86.9
"		Insulator	512×341	92.5
		U-bolt	512×341	83.8

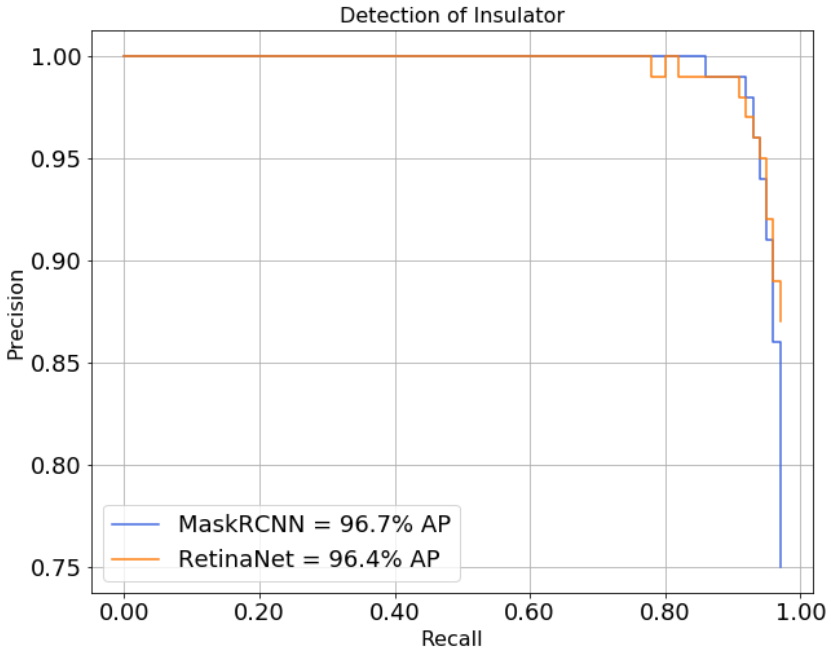


Fig. 5.6 Precision-Recall curves for the detection of insulators on high-risk towers

5.3.3 Single and multi-task detection

As mentioned earlier, U-bolts are classified based on the assets they support. Figure 5.4, (a) and (b) are examples of different U-bolt regions with closely fitting bounding boxes and there is no clear distinction between the objects in the images. Both objects have "U"

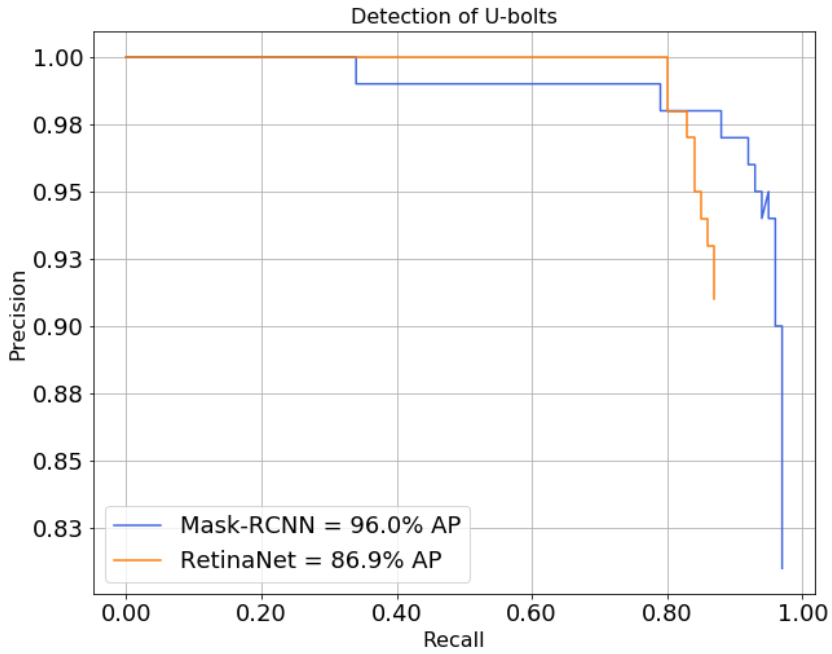


Fig. 5.7 Precision-Recall curves for the detection of U-bolts on high-risk towers

shaped structure with a link below. Figure 5.4, (c) and (d) show a different annotation option that identifies two classes of U-bolts from less closely fitted bounding box version. Insulator U-bolts and earth-wire U-bolts were detected by expanding the bounding boxes to build in some contextual details. For completeness, bounding boxes for insulators were also expanded to introduce additional contextual information, which might be useful for detection. RetinaNets were trained on the expanded bounding boxes for single and multi-task object detection. Single task detection models individual object classes separately and multi-task combines multiple classes simultaneously.

As shown in Table 5.5, there are 1,944 insulators, 1,038 insulator U-bolts and 479 earth wire U-bolts in the training set and 489, 234 and 103 insulators, insulator U-bolts and earth wire U-bolts in the testing set, respectively. As shown in Table 5.6, training the detector with multiple object classes improved performance. For example, the detection of earth-wire U-bolts increased from 90.7% to 95.1%. Similarly, insulator U-bolts detection from 96.1%

Table 5.5 Sub-image regions with more context annotations

	Towers	Images	Insulators	Insulator U-bolts	Earth wire U-bolts
Training	598	1,466	1,944	1,038	479
Testing	150	364	489	234	103
Overall	746	1,830	2,433	1,272	582

Table 5.6 Single and multi-task detection

Components	1-class AP_{50}	3-classes AP_{50}
Earth-wire U-bolts	90.7	95.1
Insulator U-bolts	96.1	97.9
Insulators	95.3	96.7

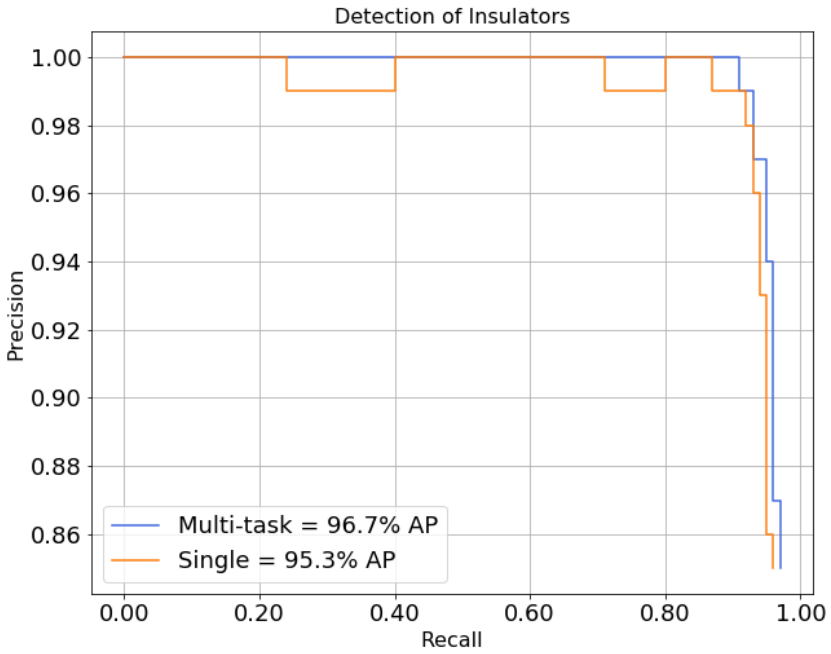


Fig. 5.8 Precision-Recall curves for the detection of insulators - single vs. multi-task

to 97.9%. Figs. 5.8 and 5.9 show the precision-recall curves comparing single and multi-task detectors for insulators and insulator U-bolts, respectively.

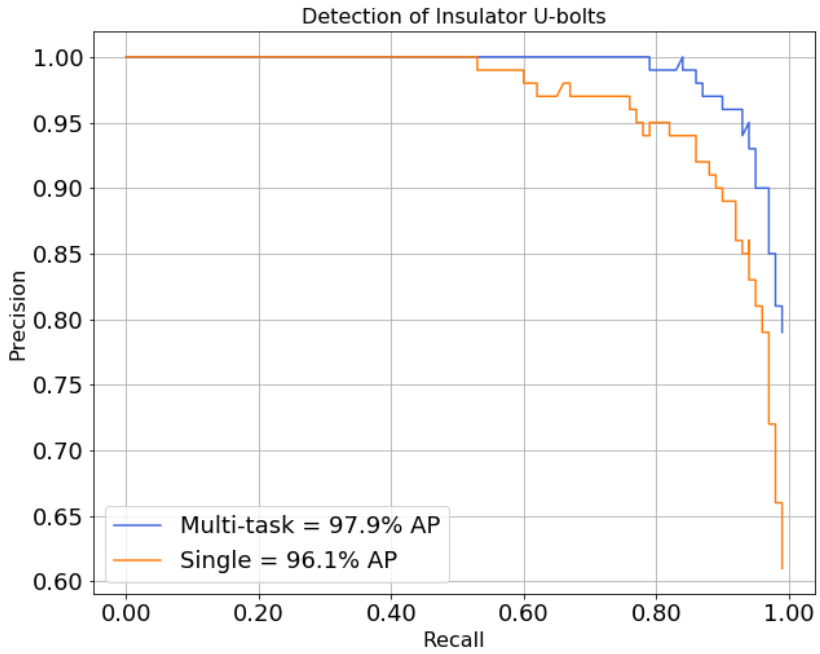
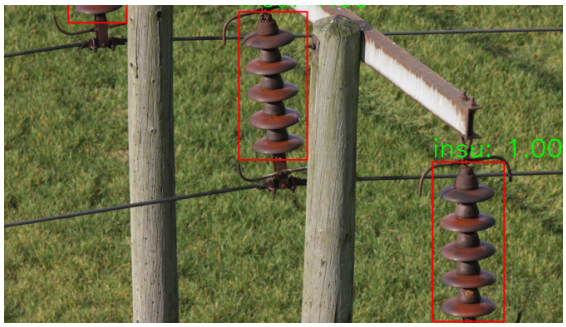


Fig. 5.9 Precision-Recall curves for the detection of insulator U-bolts - single vs. multi-task

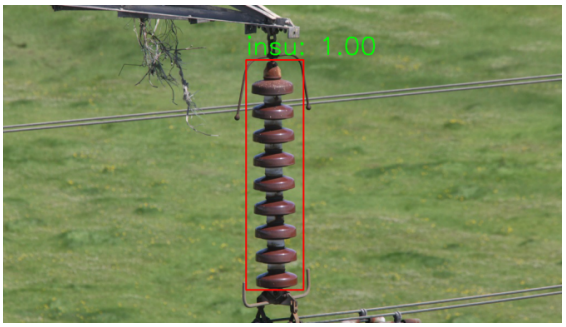
5.4 Summary

Insulators and U-bolts were detected from real-world data. The data set presents a wide variety of component types, failure modes and views, covering the complexity of aerial images along electricity networks (e.g., roof tops, vegetation, etc.). The diversity in component features and image scenes is beneficial for generalisation. Two object detection models that have not been compared for electricity network components identification were tested. Mask R-CNN and RetinaNet were extensively evaluated for the detection of varied insulator types and U-bolts, comparing the effects of (1) occlusion (2) input size, and (3) the effect of multi-task detections.

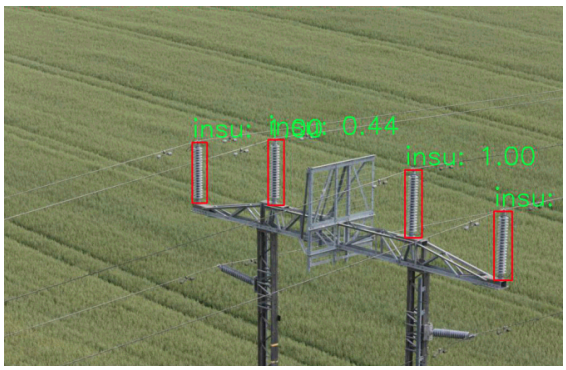
Fig. 5.10 show examples of towers with overlaid bounding boxes of different insulator types i.e., porcelain, glass, and composite types, detected simultaneously. To investigate occlusion and the effect of object scale on detection, insulators were detected from randomly



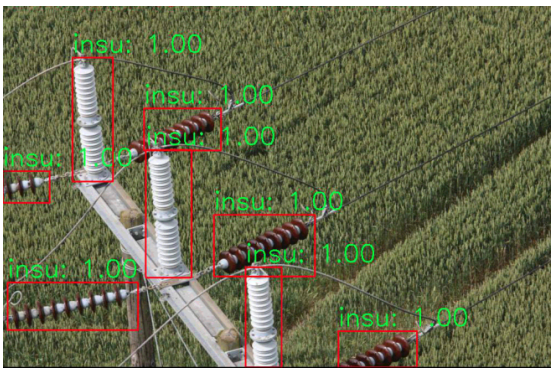
(a) Porcelain type (suspension)



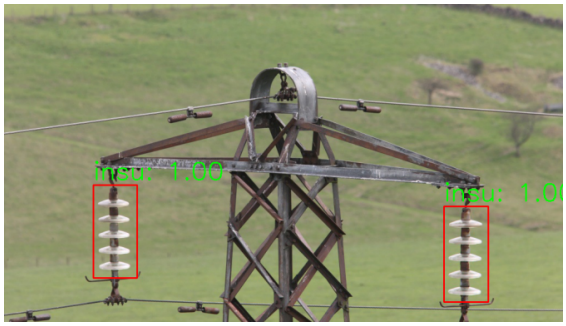
(b) Porcelain type (suspension)



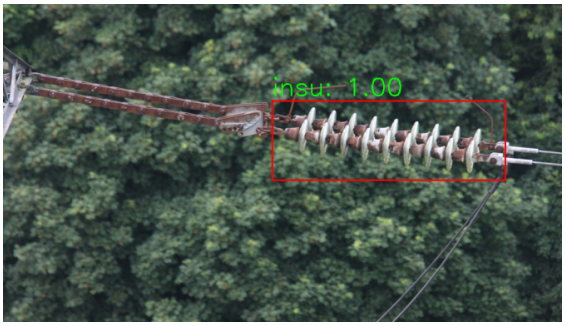
(c) Composite type



(d) Mixed of porcelain and composite



(e) Glass type (suspension)



(f) Glass type (tension)

Fig. 5.10 Insulator detection with bounding boxes overlaid

selected images. Mask R-CNN detected 1,102 regions and 25% of the detected instances were false positives. Some insulators were completely missed, and these tend to be instances that were heavily occluded by tower structure and image boundaries. It is likely that the occluded instances would be detected from other viewpoints. As shown in Table 5.2, discarding the 115 heavily occluded instances improved the performance by over 6%. Scaling input resolution downwards by half degraded the performance by about 3%. Mask R-CNN and

RetinaNet were tested on selected images of high-risk towers. Mask R-CNN recorded 96.7% AP on insulator detection and that is 7% better than its performance using images sampled at random. The images referenced from free-text were close-shot images with better resolved object regions than in randomly sampled images. The network performance shows that the scale of objects in the image is important for the task. This requirement is consistent with the report in [49], which recommended close-up images for insulator detection. It seems that when towers are suspected to be at risk, they were photographed more closely than those at normal operating conditions.

Mask R-CNN is more robust for detecting small regions like U bolts. Two-stage models are known to achieve better performance over small objects than one-stage detectors [9]. For bigger regions like insulator detection, marginal difference was recorded between these model types. At inference, RetinaNet detected objects 3x faster than Mask R-CNN.

Ling et al. [61] detected glass insulator. Porcelain and composite insulators were detected in [24, 74]. Meanwhile, [84, 103] noted that several insulators with diverse shapes, length and orientations may support a single transmission tower. A wider range of insulator types were detected in this Chapter and is useful for automatic tower assessment. Four insulator types i.e., porcelain, ceramic, glass, and composite were detected simultaneously. The results also show that multi-task detection improved performance and should be used where possible. Actual assessment of towers based on the automatically detected components follows in subsequent Chapters.

Chapter 6

Tower Rating Based on Detected Inspection Parameters

6.1 Introduction

The two major building blocks of our tower inspection pipeline are instance detection (e.g., insulator and U-bolt detection described in Chapter 5) and tower condition classification. A block diagram illustrating the relationship between the sub-systems is shown in Fig. 1.3. This Chapter describes the different approaches employed for the tower condition classification sub-system. This involves the classification of towers based on local regions of its images (i.e., the extracted regions of interest). To achieve this, multiple images of a tower were passed through a detection subsystem to extract insulators and U-bolts. Examples of insulators and U-bolts regions on tower images are shown in Figs. 5.1 and 5.2, respectively.

Importantly, our method for tower condition classification does not require explicit labelling of the multiple instances of either insulators or U-bolts. Two methods namely instance and multiple instance learning (MIL) based classifiers were explored for this problem. In this application, a tower was presented as a bag of images but with a label Y (tower condition rating). Varying numbers of images were taken from each tower such that $B =$

$\{X_1, X_2, \cdots\}$, B represents a bag and X represents images. Object instances are the condition parameters such as insulators and U-bolts detected in images.

6.2 Instance-based classification

For instance-based method, a tower's label was transferred to all the detected sub-image regions. Image instance classifiers namely ResNet50V2 [36] and EfficientNetB0 [114] were compared for instance classification. These are standard deep learning models available in Keras for image classification. In Chapter 4, a ResNet50V2 was fine-tuned for the classification of tower configuration. In this Chapter, we simply changed the output parameter of these networks and trained from scratch to classify detected instances as healthy or unhealthy. In other words, the number of classes for each network was changed to 2 instead of 1,000 classes for ImageNet dataset. While there are 23M parameters in ResNet50V2, EfficientNetB0 has only 4M parameters. Stochastic gradient descent (SGD) with initial learning rate $\alpha_i = 0.001$ was applied. The learning rate was then scaled based on Eqn. 6.1.

$$\alpha = \alpha_i(1 - \frac{e}{e_{max}}) \tag{6.1}$$

where α_i is the initial learning rate, e is the current epoch and e_{max} is the maximum number of epochs. At test time, tower was classified by aggregating the class probability scores of all instances as shown in Fig. 6.2.

6.3 MIL-based classification

Our MIL network makes use of EfficientNetB0 as the base network for feature extraction. The base network was chosen because it is optimised in terms of network depth, width, and resolution. It means that the input size determines the scale of the network and hence

Table 6.1 EfficientNetB0 Architecture

Stage	Operator	Resolution	Channels	Layers
1	Conv3x3	224 x 224	32	1
2	MBConv1, k3x3	112 x 112	16	1
3	MBConv6, k3x3	112 x 112	24	2
4	MBConv6, k5x5	56 x 56	40	2
5	MBConv6, k3x3	28 x 28	80	3
6	MBConv6, k5x5	14 x 14	112	3
7	MBConv6, k5x5	14 x 14	192	4
8	MBConv6, k3x3	7 x 7	320	1
9	Conv1x1 & Pooling & FC	7 x 7	1280	1

the numbers of parameters. Table 6.1 shows the architecture of EfficientNetB0 [114]. The network was implemented in 9 stages. Stage 9 has a 1×1 conv layer, pooling and a dense layer. We flattened the conv. layer and fed that into a new fully connected (FC) layer with 64-nodes. Two dropouts were applied immediately after the FC layer. Dropout is a technique to reduce the complexity of a network by randomly dropping some nodes of a layer and hopefully minimise over-fitting. A second FC with 32-nodes was added and followed by a set of dropouts. Additional regularisation using L2 (weight decay) was applied. Ilse et al. [44] presented attention-based MIL pooling, a weighted average of instances with weights determined by a neural network. We used the attention-based MIL as the final layer. Fig. 6.1 shows a deep multiple instance learning network with 8M parameters and trainable end-to-end.

6.4 Training and evaluation

MIL learns bags of representations as opposed to learning individual instances. Algorithm 3 shows the steps for the MIL classification of towers based on all detected instances. Training and evaluation were carried out on RTX2080 machine with a 10GB GPU memory. Although this hardware specification is adequate for instance classification as shown in Algorithm 2, it

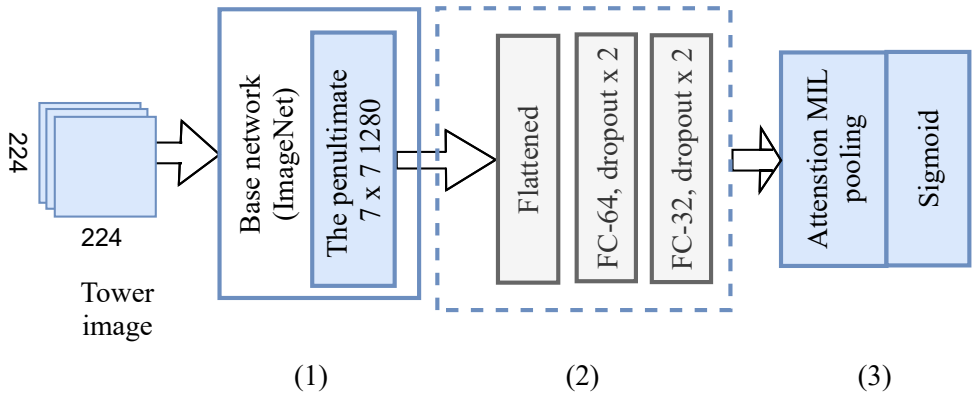


Fig. 6.1 MIL-classifier (1) Feature extraction using EfficientNet-B0 [114], (2) Added FC layers with dropouts, (3) MIL pooling as a weighted average of instances

is challenging for MIL-based classification. It is difficult to know in advance how much GPU memory will be required for an experiment. The amount of memory available determines for example, the batch size for image classification task. Our instance classifiers were trained using a batch size of 8 images with dimensions 224×224 pixels. Training a MIL-based network using all detected instances as shown in Algorithm 3 is more computationally demanding. A bag represents a tower, and the distribution of instances vary across towers. We could train our MIL network (Fig. 6.1) using bags of 50 images without out-of-memory (OOM) error. A MIL using bags of 50 selected instances is known as MIL-50. To utilise all detected objects, which in most cases are 100s, tower bags were split into sub-bags as shown in Fig. 6.3. Manivannan et al. [70] supports the splitting of a large input image into subsets for a classification task. Subsets of a fundus camera image are classified and aggregated to determine the image label. Here, experiments were carried out using sub-bags of 8 instances (MIL-8) and 16 instances (MIL-16). All tower sub-bags remain within either the training or testing split to avoid bias. During training using sub-bags, a tower label Y was transferred to all sub-bags generated from it. Table 6.2 shows a summary of selected network training parameters.

The classifiers were evaluated using 10-fold cross-validation. There are several methods used for aggregating bag representation including mean pooling, max pooling, and log-sum-

Table 6.2 Parameter selection and the values used for tower condition classifiers. The initial learning rate was scaled by Eq. 6.1

Parameter	Instance-based	MIL-8	MIL-16	MIL
Size of input image	224×224	☑	☑	☑
Maximum number of epochs	25	☑	☑	☑
Optimisation function	SGD	☑	☑	☑
Initial learning rate	0.001	☑	☑	☑
Number of instances per batch	8	8	16	50
Evaluation	10-folds cross-validation	☑	☑	☑

Algorithm 2: Instance-based tower classification

Data: *tower*
Result: *ConditionRating*
$X \leftarrow \{\}$;
$N \leftarrow len(X)$; /* Number of instances for the tower */
$cr \leftarrow \{\}$;
while $img \in tower$ **do**
  $image \leftarrow imageRead(img)$; /* Read an image of $N \times N$ pixels */
  $X \leftarrow \delta(image)$; /* X = bag of sub-image regions of interest */
  while $inst \in X$ **do**
    $y \leftarrow f(inst)$; /* Classify instance as healthy/unhealthy */
    $pos \leftarrow y(unhealthy)$; /* Score corresponding to unhealthy */
    $cr \leftarrow pos$; /* Update list of instance ratings */
  $ConditionRating \leftarrow \frac{\Sigma(cr)}{N}$; /* Aggregate instance ratings */

exp pooling [44, 122]. In these experiments, tower label, Y was determined using either by averaging predictions or by using that max of the predictions. Averaging predictions be it for instance-based or sub-bags is likely to work well where conditions of instances on a tower are similar. The max operator selects the highest predicted instance or sub-bag probability as the tower label. Performance was measured using the receiver operating characteristic (ROC) curve. The ROC shows the relationship between sensitivity and specificity over different decision thresholds. Sensitivity is the ability of the instrument to correctly predict positive targets. Specificity on the other hand is the probability of a negative target to have a negative test result. The resulting ROC curve shows the different operating points for the

Algorithm 3: MIL-based tower classification

Data: *tower*
Result: *ConditionRating*

$X \leftarrow \{\}$;
$cr \leftarrow \{\}$;
while $img \in tower$ **do**
    $image \leftarrow imageRead(img)$; /* Read an image of $N \times N$ pixels */
    $X \leftarrow \delta(image)$; /* X = bag of sub-image regions of interest */
    if $len(X) \leq 50$ **then**
        $ConditionRating \leftarrow f(X)$; /* Rate bag as (healthy/unhealthy) */
    else
        $subBags \leftarrow \frac{X}{n}$; /* n is the sub-bag size */
        while $subBag \in subBags$ **do**
            $y \leftarrow f(subBag)$; /* Classify sub-bag (healthy/unhealthy) */
            $pos \leftarrow y(unhealthy)$; /* Score corresponding to unhealthy */
            $cr \leftarrow pos$; /* Update list of sub-bag ratings */
        $ConditionRating \leftarrow \frac{\Sigma(cr)}{len(subBags)}$; /* Average sub-bags preds. */

instrument. In this experiment, the best operating point was determined using Youden's index, J (Eqn. 6.2) [113]. Although Youden's index is relatively simple to apply, it assumes equal costs for both false-positives and false-negative prediction and does not take into consideration the effect of the prevalence of individual class [113].

$$\begin{aligned} J &= Sensitivity + Specificity - 1 \\ &= \text{True-positive rate} + (1 - \text{False-positive rate}) - 1 \\ &= \text{True-positive rate} - \text{False-positive rate} \end{aligned} \tag{6.2}$$

6.5 Testing

The remaining part of this chapter presents the evaluation of the proposed methods. Two inspection parameters were selected for assessment namely insulators mechanical failure

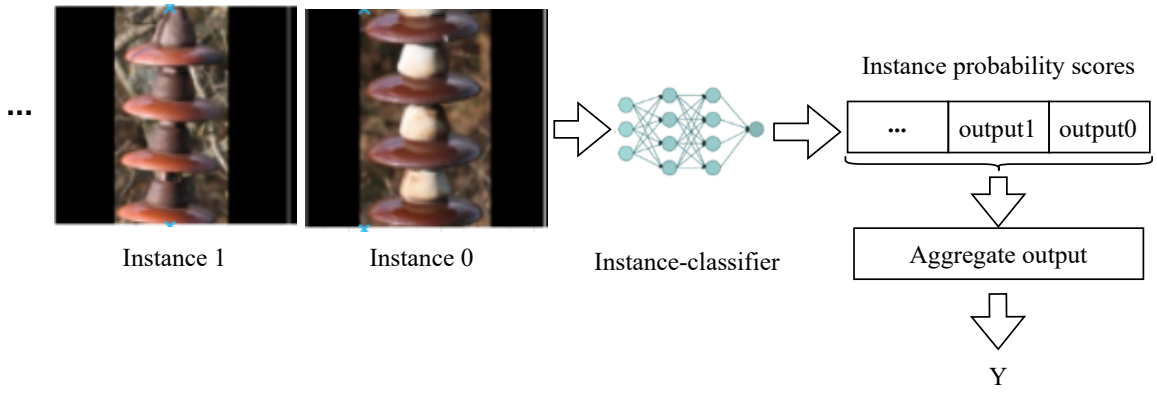


Fig. 6.2 Instance-based tower classification

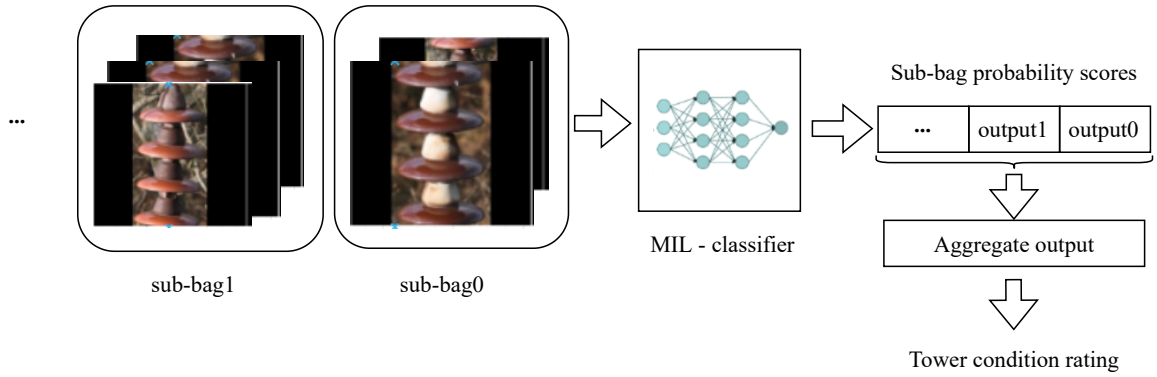


Fig. 6.3 Sub-bags for MIL with limited GPU memory

mode and U-bolts. The classifiers were compared using area under the ROC curve, AUC and confusion matrices computed at optimal thresholds.

6.5.1 Insulator Failure Modes

Insulators play a critical role in OHL towers. Chapter 5 demonstrated the detection of insulators and U-bolts from aerial images with reasonable recall and precision. The detection of these assets is an important step for the automation of routine tower assessment. Tower ratings are measured by the failure of individual components that they support. There are two categories of faults on insulators that are visible in an image upon which towers are rated. The faults can be categorised into electrical and mechanical failure modes. Electrical failure

modes are faults that reduce or undermine the electrical properties of the device leading to energy losses, etc. Li et al. [56], Sampedro et al. [103] highlighted regions of insulators with cracks or chips on insulator sheds or missing sheds in glass insulator type. Examples of this failure modes are shown in Fig. 6.4. Detecting sub-insulator region like that is important for the automated inspection of towers but the industry want the classification of this failure at tower level.

In contrast to detecting missing and chipped sheds on insulators, mechanical rust involves corrosion (rust) of the entire insulating string as shown in Fig. 6.4 (4 and 5). Corrosion can be localised around the metal cap or pin and can also affect the entire insulator string. Assessing the level of rust or corrosion of insulators from aerial images presents significant challenges arising from scale, illumination changes and sub-types of the components (porcelain, glass, composite, and ceramic), which must be assessed simultaneously. Jalil et al. [47] reported a DL method that detected insulator instances as rusty (or not). A total 160 insulators were extracted from 132 images and augmented by performing series of offline image transformations of the original insulators [47]. Image augmentation is particularly useful for challenges like lighting, occlusion, scale, background and can prevent over-fitting the model but will not cause a model to generalise on a target they have not seen previously [110]. Detection of insulator instances as presented in [47, 56, 103] would need fine-grained labelling of the assets and that is costly. An alternative to instance condition classification is to deal with tower level condition rating, which might be more useful.

The criteria for measuring the extent of rust of a tower based on insulators was presented in Chapter 3. A tower was rated as either CR-1, CR-2, CR-3, CR-4, or CR-5 depending on the severity of rust on the insulators, with CR-1 meaning as new and CR-5 denoting a high-risk situation. It is important to emphasise that tower condition ratings are at the tower level and not at component (i.e., insulator) level. It is the tower that is assigned a condition rating; no condition labels were available in the database for the individual components on

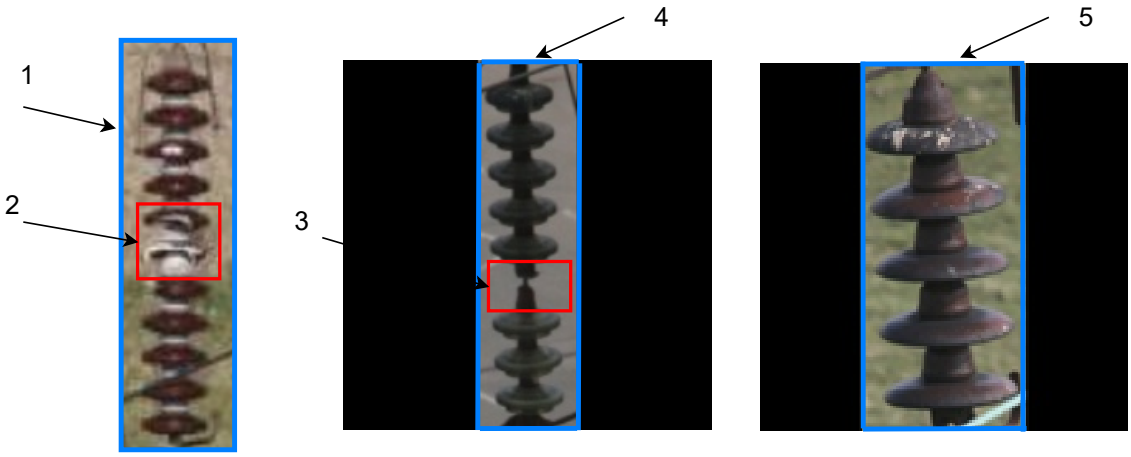


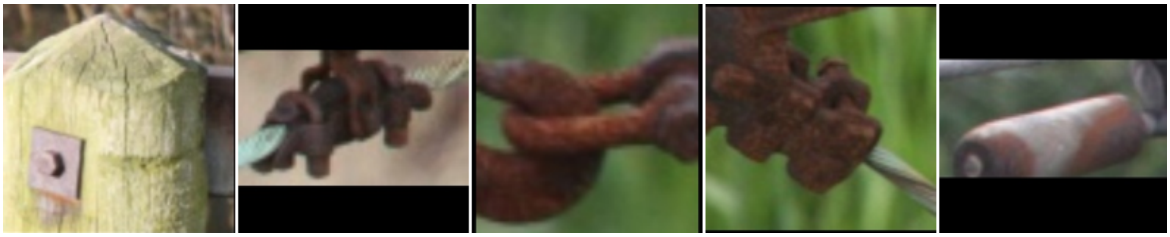
Fig. 6.4 Suspension insulator with examples of failure modes: (1) Insulator string. (2) Chips on insulator sheds. (3) Missing shed. (4 and 5) Severely rusty insulator string.

which the tower ratings were based. The critical decision threshold for this failure mode lies between CR-3 and CR-4. Towers given ratings of CR-4 or CR-5 will be scheduled for immediate refurbishment. Therefore, a binary classifier was developed to discriminate between *healthy* towers (CR-1, CR-2, CR-3) and *unhealthy* towers (CR-4, CR-5). Unhealthy towers are regarded as the positive class for the binary classification task.

As shown in Table 6.3, there are 5,301 towers that do not overlap with those used for training our insulator detectors. This study is interested in insulators and because insulators are found around the cross-arm of towers, 3 image categories were used in this experiment. Specifically, images labelled *insulator top*, *insulator middle* and *insulator bottom* (see Table 3.5). Both S-type and T-type tower types were selected. There is an average of 25 images from each tower. For each condition rating, 200 towers were sampled and classified. Recall that towers rated CR-1, CR-2 and CR-3 are at normal operating conditions while a those rated CR-4 and CR-5 would be scheduled for refurbishment. Insulators were automatically extracted from the multiple images of each tower and examples of extracted regions are shown in rows 1 and 2 of Fig. 6.5. The number of insulator instances across towers vary depending on the type of tower and the number of images provided. Fig. 6.6 shows the distribution for 1,000 towers in the data sample and has mean and standard deviation of

(a) A mixture of glass, ceramic, and porcelain insulator types extracted from a healthy tower

(b) A mixture of corroded and non-corroded insulator instances from an unhealthy tower



(c) Typical false-positive detections

Fig. 6.5 Examples of automatically extracted sub-image regions (insulator instances).

82.2 and 75.9, respectively. All the regions extracted were used including glass, porcelain, composite, and ceramic insulator types. Some towers have a mixture of insulator types. The extracted regions also include false-positive detection from each tower and examples are shown in row 3 of Fig. 6.5. Each detected region was then cropped and resized to input size of 224×224 pixels.

A tower normally would be supported using similar component types and these would age uniformly. However, new components and sometimes of different designs are introduced within a tower's life cycle. The variability of inspection component such as different insulator types could impact the performance of automatic tower condition assessment negatively.

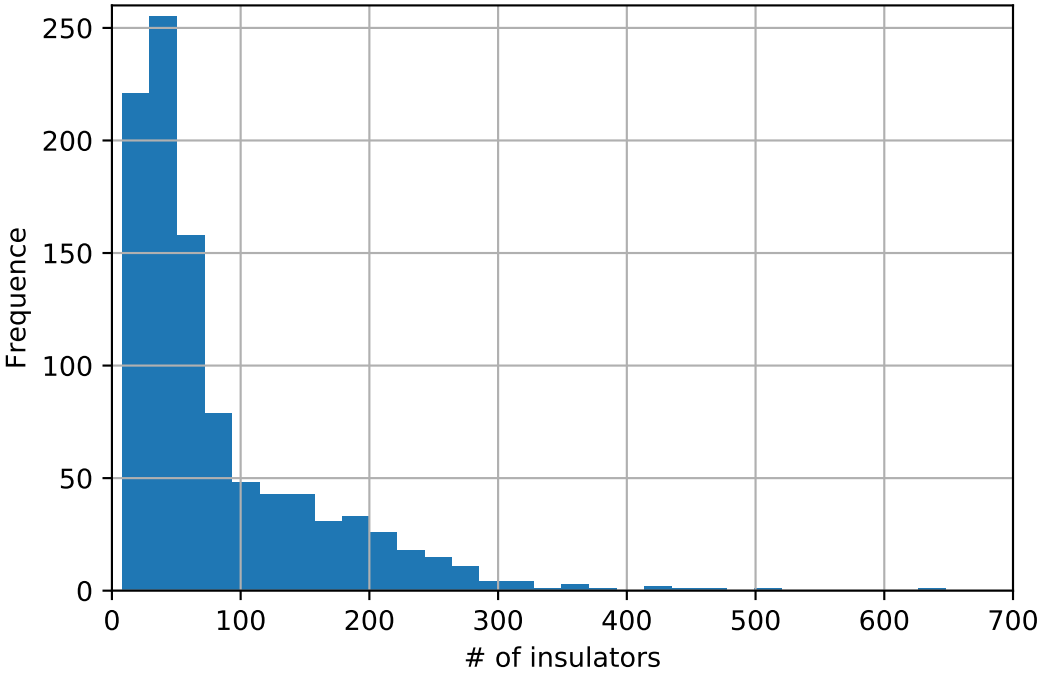


Fig. 6.6 Distribution of detected insulators.

Table 6.3 Distribution of towers, images and extracted 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

Instance-based classification

Fig. 6.7 shows the ROC curves for EfficientNet, achieving 0.94 AUC and 0.90 AUC on averaging and by taking the max of instance class predictions, respectively. Also, Fig. 6.8 shows the ROC of ResNet, achieving 0.90 AUC and 0.81 AUC based on averaging and max operators on instance predictions, respectively.

The confusion matrices shown in Table 6.5 were computed at optimum thresholds. The optimum threshold values for EfficientNet were 0.4182 and 0.9426 for mean and max,

Table 6.4 Summary of performance: Tower condition classification

Method	Network Param.	Insulator-based (AUC) MEAN	MAX
EfficientNet	4M	0.94	0.90
ResNet	23M	0.90	0.81
MIL-8	8M	0.90	0.88
MIL-16	8M	0.92	0.91

Table 6.5 Confusion matrix: Instance-based tower classification

		ResNet		EfficientNet	
		Healthy	Unhealthy	Healthy	Unhealthy
MEAN	Healthy	516	84	506	94
	Unhealthy	75	325	42	358
MAX	Healthy	460	140	487	113
	Unhealthy	110	290	46	354

respectively. Whereas, ResNet had 0.4654 and 0.9999 optimum threshold values for mean and max, respectively. ResNet followed by averaging instance predictions classified 841 towers correctly and misclassified 159 towers. By averaging instance predictions, EfficientNet classified 864 towers correctly and missed 136 towers. EfficientNet predictions followed by the maximum of instance predictions predicted 841 towers correctly and missed 159 towers. ResNet with max classified predicted 750 towers correctly and misclassified 250 towers. Table 6.6 is an expanded confusion matrix of EfficientNet-based classifier. It shows how the classification results compare to the ground-truth. Condition ratings CR-1, CR-2, CR-3 are healthy towers whereas CR-4, CR-5 are unhealthy (positive class).

MIL-based classification

As shown in Fig. 6.3, multiple sub-bags of each tower were classified to utilise all insulator instances. Fig. 6.9 shows the ROC curves for MIL-8 based classifier, comparing averaging, and the max of sub-bag predictions. Averaging sub-bags probability scores achieved 0.90

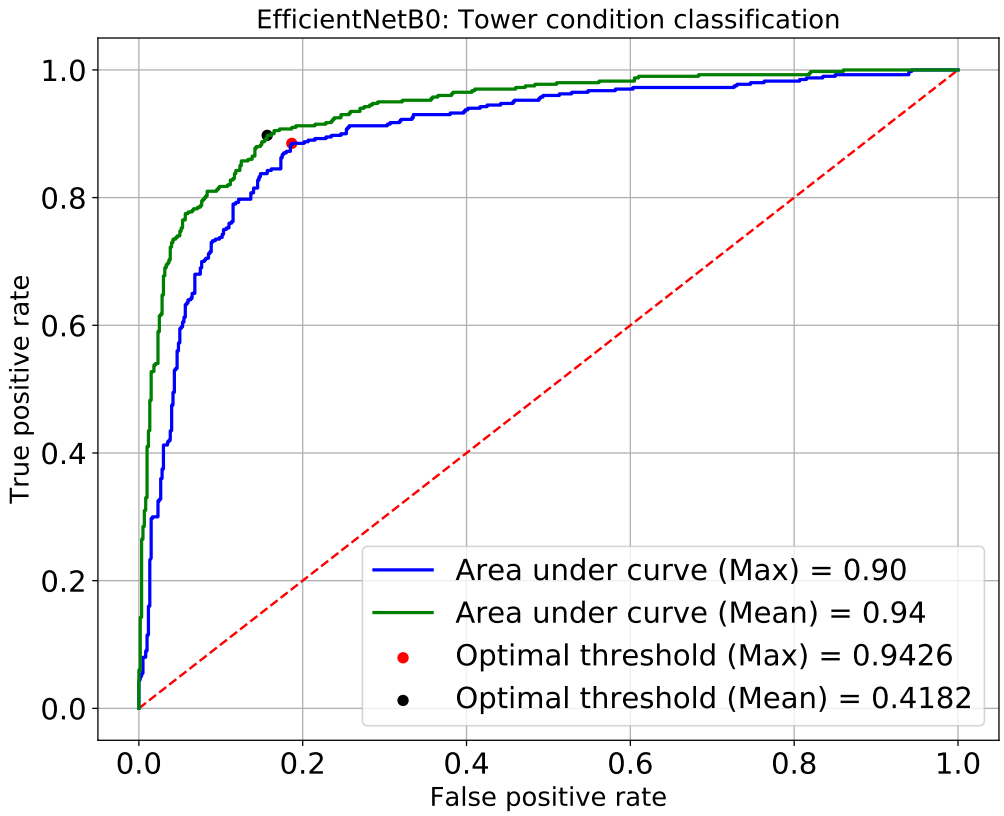


Fig. 6.7 ROC curves for tower condition classification: EfficientNet instance-based

Table 6.6 Expanded confusion matrix using EfficientNet

		MEAN		MAX	
Class	CR	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	195	5	186	14
	2	169	31	165	35
	3	142	58	136	64
Unhealthy	4	30	170	27	173
	5	12	188	19	181

AUC while the max produced 0.88 AUC. Table 6.8 shows the confusion matrix computed at optimum thresholds of 0.5752 and 0.9999 for mean and max, respectively. Averaging of sub-bags (8 instances) classified 823 towers correctly and missed 177 towers. The classifier

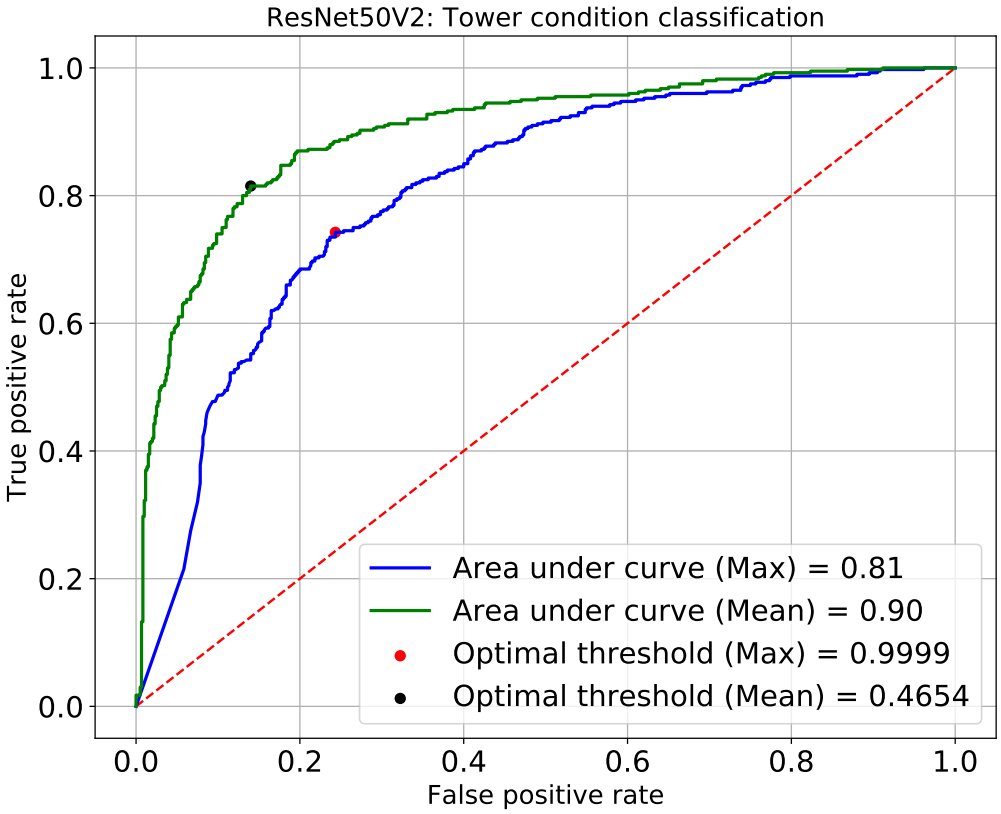


Fig. 6.8 ROC curves for tower condition classification: ResNet instance-based

Table 6.7 Expanded confusion matrix using ResNet

		MEAN		MAX	
Class	CR	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	187	13	165	35
	2	175	25	151	49
	3	154	46	144	56
Unhealthy	4	47	153	58	142
	5	28	172	52	148

based on taking the max of sub-bags predictions resulted to correct classification of 812 towers while 188 towers were wrongly classified.

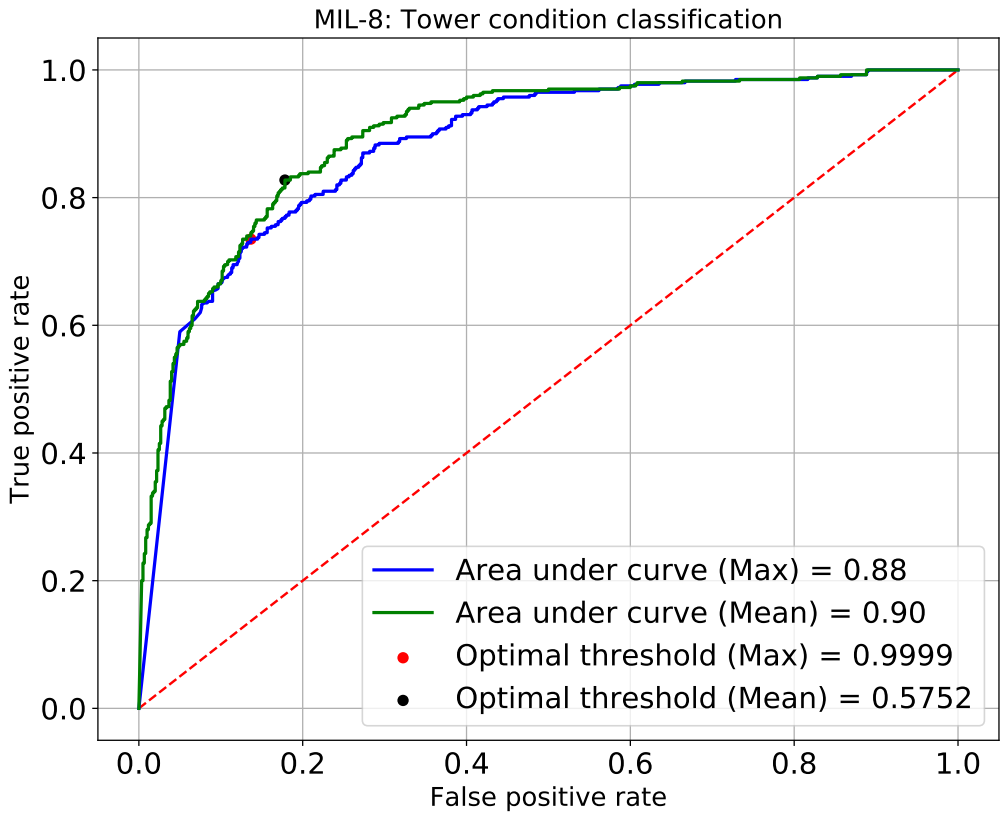


Fig. 6.9 ROC curves for tower condition classification: MIL-8

As shown in Table 6.4, increasing the number of instances contained in each sub-bag from 8 to 16 instances improved AUC values from 0.90 to 0.92 on mean aggregation. Sub-bags with more instances are more likely to have positive examples. The ROC curves of the classifier are shown in Fig. 6.10. Confusion matrices shown in Table 6.8 was computed at 0.4999 and 0.9357 optimum threshold values for mean and max respectively. By averaging sub-bag (16 instances), the classifier predicted 838 towers correctly and missed 162 towers. Sub-bag predictions followed by max operation classified 821 towers correctly and missed 179 towers. Further, tower classification results were compared with the original tower labels in Tables 6.9 and 6.10 for MIL-8 and MIL-16, respectively.

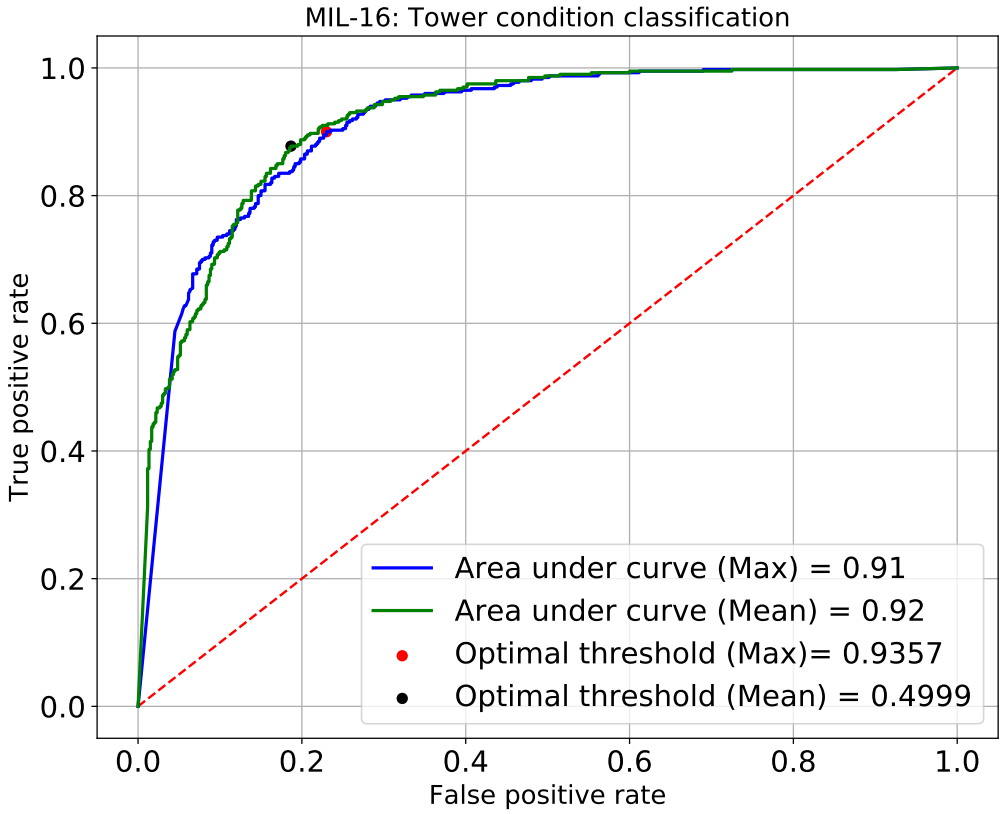


Fig. 6.10 ROC curves for tower condition classification: MIL-16

Table 6.8 Confusion matrix: MIL-based tower classification

		MIL-8		MIL-16	
		Healthy	Unhealthy	Healthy	Unhealthy
MEAN	Healthy	493	107	488	112
	Unhealthy	70	330	50	350
MAX	Healthy	520	80	462	138
	Unhealthy	109	291	41	359

6.5.2 U-bolt Failure Modes

U-bolts are critical asset on a tower as well and are visible on suspension type of towers. Chapter 4 presented the classification of towers into suspension and tension, which is believed to be useful for the detection of components. Electrical cables are attached to suspension

Table 6.9 Expanded confusion matrix of MIL-8

		MEAN		MAX	
Class	CR	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	190	10	190	10
	2	160	40	180	20
	3	143	57	150	50
Unhealthy	4	43	157	63	137
	5	27	173	46	154

Table 6.10 Expanded confusion matrix of MIL-16

		MEAN		MAX	
Class	CR	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	190	10	183	17
	2	168	32	163	37
	3	130	70	116	84
Unhealthy	4	35	165	29	171
	5	15	185	12	188

towers using a system of insulator and U-bolt (see Fig. 6.11). Insulator U-bolts bear the load of cables suspended by them. Due to corrosion and mechanical erosion at the link point, the component fails. Electricity network operators assess towers by this failure mode. A failure at this region would affect the structural integrity of the tower and network.

As described in Chapter 3, tower conditions based on their U-bolts or insulators were labelled separately for right and left circuits. Typically, the number of circuits on a line determines the number of insulating points required on a tower. We have single and double circuit towers in our datasets and to utilise all in our experiments, we combined the right and left circuits condition ratings by computing the maximum of both circuits. It seemed reasonable to think that if the left circuit of a tower is at high risk, the whole tower is at risk and vice versa. This Section demonstrates the classification of towers using automatically detected insulator U-bolts. Fig. 6.11 shows a tower image with a bounding box overlaid on insulator U-bolt region.

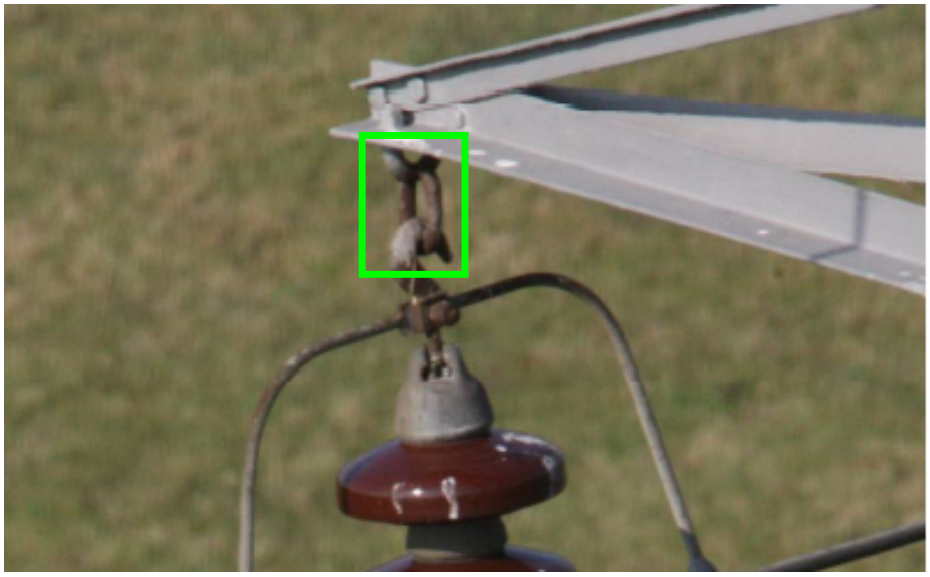


Fig. 6.11 Insulator U-bolt

Table 6.11 Distribution of towers and insulator U-bolts (sample)

Class	Overall Towers	Sample Towers	Sample Images	Extracted U-bolts
Healthy	3,324	320	8,663	9,168
Unhealthy	323	320	7,061	10,083
Total	3,647	640	15,724	19,251

Furthermore, U-bolts are visible on suspension (S-type) towers. All S-types towers in the data set were used including those used previously for training detectors. We re-used some towers because of the small number of high-risk (unhealthy) class. As shown in Table 6.11, there are a total of 3,647 S-type towers. 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. A data set was created for this purpose using a sample of the suspension (S-type) towers. Only 323 of the 3,647 S-type towers are categorised as unhealthy so to create a balanced data set, 320 of healthy and unhealthy towers were sampled. 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

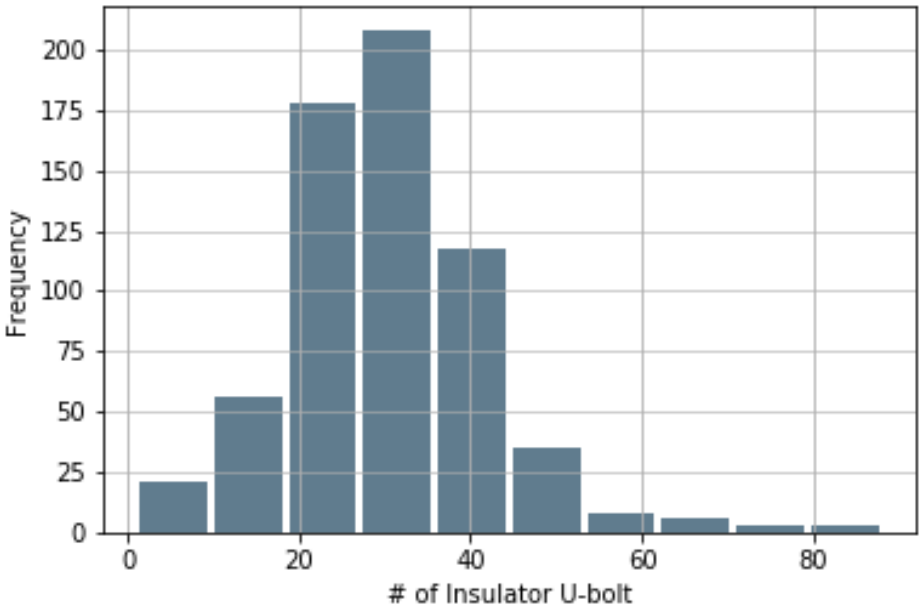


Fig. 6.12 Distribution of number of detected insulator U-bolts per tower

Table 6.12 Distribution of insulator U-bolt (overall)

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

instances (sub-images regions) of insulator U-bolts were automatically detected. Examples of U-bolt regions were presented in Chapter 5. Table 6.11 gives a detailed breakdown of the dataset and the distribution of extracted insulator U-bolts is shown in Fig. 6.12. There is an average of 30 U-bolts detected per tower with 24 towers having between 51 and 88 instances (upper boundary). The number of detected instances depends on the number of photographs provided for a tower. All automatically detected instances were used in what follows, including false positive detection. Each detected instance sub-image was scaled, cropped and zero-padded to 224×224 pixels.

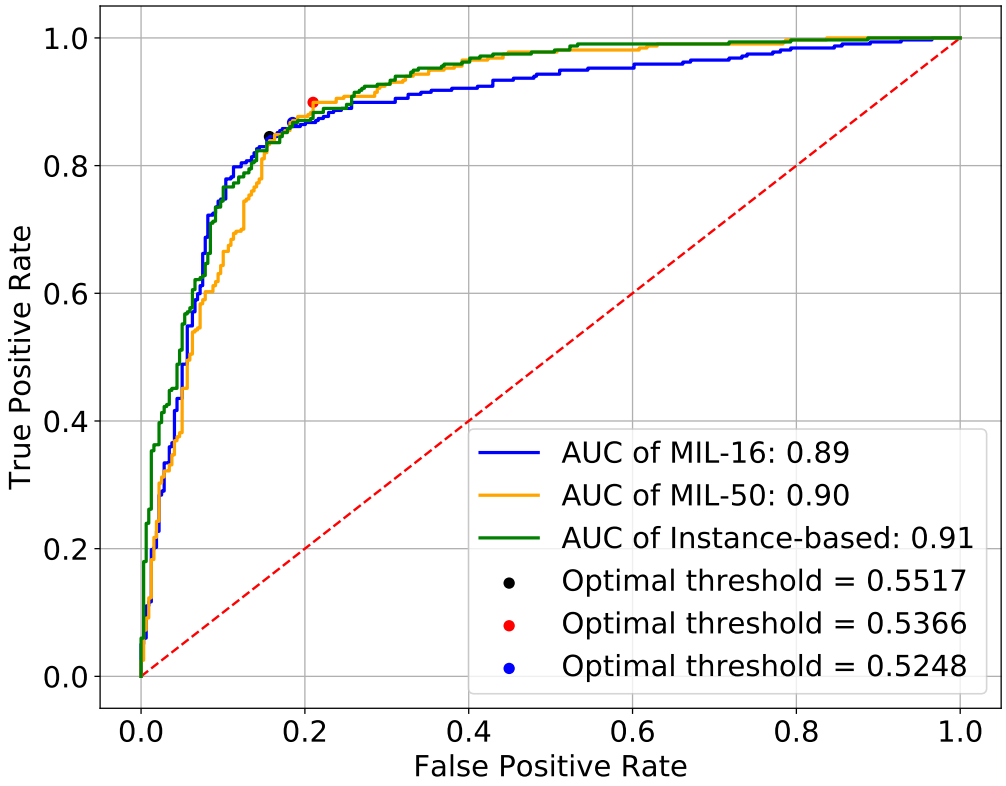


Fig. 6.13 ROC curves of Instance-based, MIL-16 and MIL-50

MIL-based classification

A MIL was tested by using only selected 50 instances from each tower. The selection was based on detection probability scores (descending order) of instances. Fig. 6.13 shows ROC curves for the MIL-based tower classification with sub-bag size of up to 16 insulator U-bolts (MIL-16) achieving an AUC of 0.89. At optimal threshold of 0.5517, it classified 537 towers correctly and misclassified 99 towers. Using only selected 50 instances (MIL-50) has an AUC of 0.90 compared to 0.89 from using sub-bags of 16 and all detected instances.

Table 6.13 Confusion matrix of tower condition classification (sample)

Class	CR	MIL-50		MIL-16		Instance-based	
		Healthy	Unhealthy	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	78	4	81	1	80	2
	2	70	10	76	4	76	4
	3	104	53	112	45	104	53
Unhealthy	4	28	218	44	202	38	208
	5	5	66	5	66	5	66

Instance-based classification

Fig. 6.13 shows the ROC curve for instance-based tower classification. The classifier achieved an AUC of 0.91. A confusion matrix computed at optimum threshold of 0.53 is shown in Table 6.13. A total of 534 towers were classified correctly and 102 towers were misclassified. On the data sample, there is no significant difference in the AUCs of instance and MIL-based classifiers.

Instance-based network was tested on all 3,647 S-type towers. As shown in Table 6.12, this data set presents a highly unbalanced distribution of 3,324 and 323 healthy and unhealthy towers, respectively. The extracted instances of the healthy class are about 10 times larger than unhealthy examples. Class weights of 1:10 were applied during training to deal with this class imbalance. Fig. 6.14 shows the ROC curves overall, comparing the mean and max aggregates of instance predictions. Mean of instance predictions achieved an AUC of 0.98 whereas the max recorded AUC of 0.94. Table 6.14 shows the confusion matrices. Table 6.15 compares tower ground truth labels i.e., CR-1, CR-2, CR-3, CR-4, CR-5 with the predictions. The confusion matrix for mean of instance predictions was generated at 0.32 optimal threshold and that for max was based on a threshold of 0.99. Averaging instance prediction classified 3,480 towers correctly and missed 155 towers in total. Instance predictions followed by a max classified a total of 3,220 towers correctly and missed 415. Mean aggregation works better overall and as shown in Table 6.15, only 4 of the 72 towers

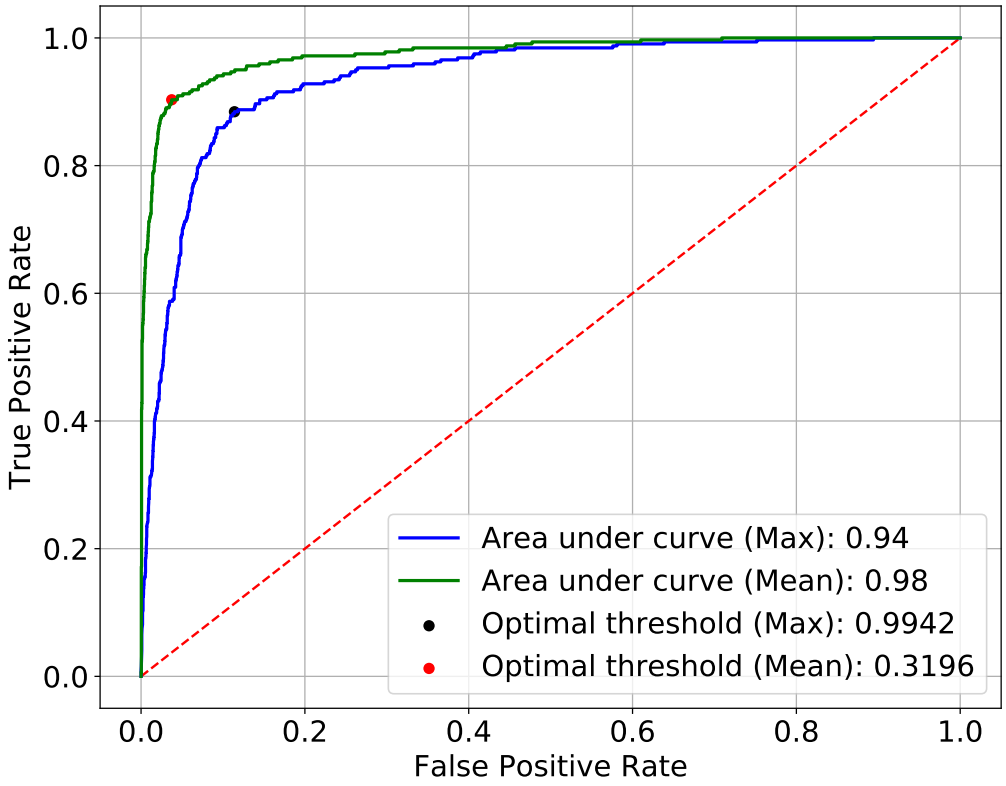


Fig. 6.14 ROC curves for tower condition classification (overall)

Table 6.14 Confusion matrix of tower condition classification (overall)

	Instance-based (Mean)		Instance-based (Max)	
	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	3,191	124	2,937	378
Unhealthy	31	289	37	283

labelled CR-5 were misclassified. Also, 1 tower was misclassified from 865 CR-1 towers. The classifiers could identify "as new" and "high-risk", which are at extreme conditions. It would have been disturbing to see serious confusion between these extreme classes. As expected, more classification errors occurred at the threshold point i.e., between CR-3 and CR-4.

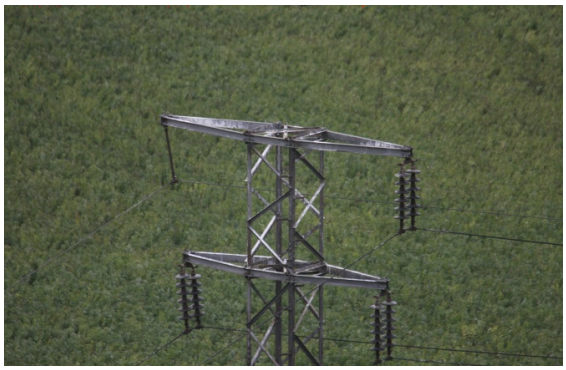
Table 6.15 Expanded confusion matrix of tower condition classification (overall)

		MEAN		MAX	
Class	CR	Healthy	Unhealthy	Healthy	Unhealthy
Healthy	1	864	1	841	24
	2	840	7	797	50
	3	1,487	116	1,299	304
Unhealthy	4	27	221	34	214
	5	4	68	3	69

Results show that although 3,647 S-type towers were considered, 3,635 were classified. Ninety six percent (96%) of the towers were correctly classified. At instance detection stage, 12 towers were excluded as not having insulator U-bolts. Most of the missed towers are low voltage towers with pin-type insulators as shown in Fig. 6.15d. This category of S-type towers is not fitted with U-bolts. So, it is a good thing that the models have learned to isolate this tower type thereby reducing false positive detection. Fig. 6.15c shows an example of a human error in which a T-type tower was wrongly labelled as S-type. It is also interesting that no sub-image region was detected on a T-type tower as they are not supported. Fig. 6.15a and Fig. 6.15b are indeed S-type towers but were missed. For some reason, insulator U-bolt's regions were occluded in all the images provided for the tower in Fig. 6.15a. The tower in Fig. 6.15b was photographed from long-range. While Fig. 6.15a and Fig. 6.15b are examples of system limitations, Fig. 6.15d and Fig. 6.15c show some data set limitations.

6.6 Summary

An electricity tower may be classified as healthy or unhealthy based on the condition parameter. In this Chapter, we experimented with insulators and U-bolts as condition parameters. These failure modes were highlighted as critical inspection parameters of OHL towers. Current industry standard relies on manual assessment of multiple insulators on tower images, which is costly and subjective. Hence, the research efforts to automate tasks



(a) Occluded insulator U-bolts

(b) Long-range tower image

(c) T-type labelled as S-type (Human error)

(d) Tower with pin insulators (not supported)

Fig. 6.15 Examples of tower excluded by the detection system

in the inspection pipeline. Two deep learning methods were tested for the classification of tower conditions from multiple insulator or U-bolt instances. The towers inspected were drawn from real-world inspection data.

Automatically extracted instances were classified by assigning tower label to each detected instance. The performance of EfficientNetB0 and ResNet50V2 were compared, achieving AUCs of 0.94 and 0.90 on averaging insulator instances predictions, respectively. As stated previously, aggregating instance class predictions would work better if the same type of component is involved and if the failure modes are evenly distributed across instances as shown in Fig. 6.16. Components installed on the same date would naturally degrade at the same rate since they would be exposed to similar loads and environmental conditions. However, it is common to find variations in conditions of insulators on the same tower due to maintenance and replacement of parts over time. As shown in Fig. 6.17, all the components

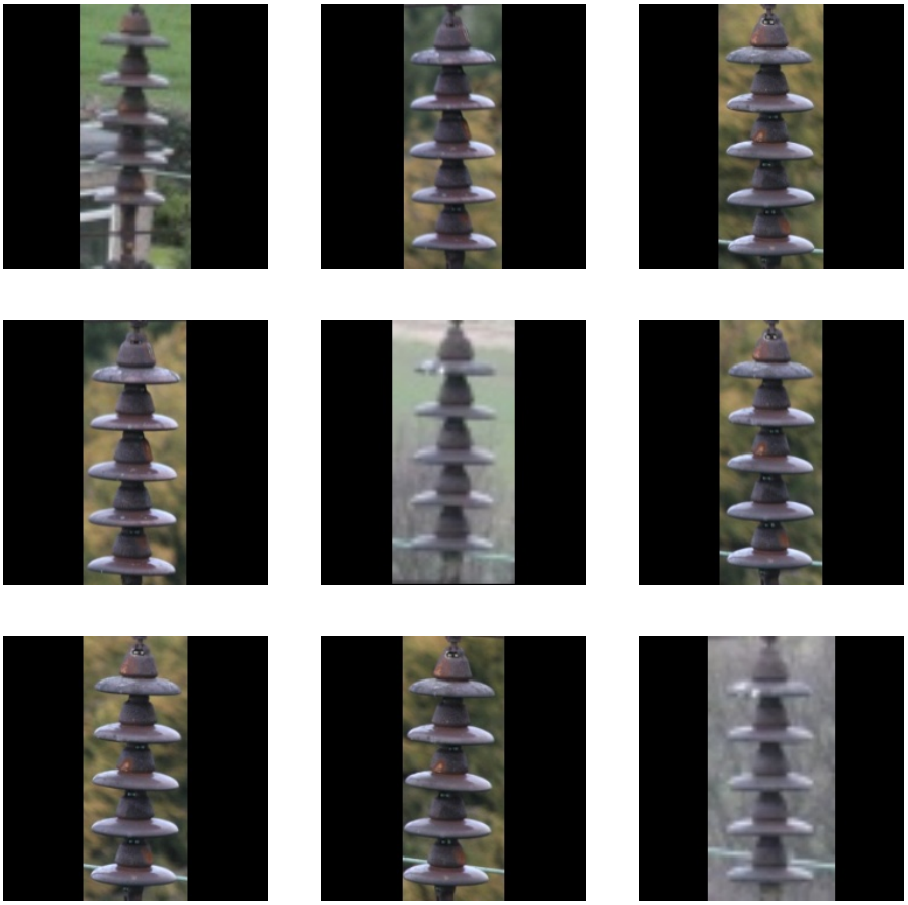


Fig. 6.16 Evenly rusty instances of porcelain insulators.

on a tower could be at different levels of degradation. Aggregating instance class predictions for such uneven distribution will be counterproductive.

A second approach involves the use of MIL for the classification of a whole tower. This method classifies a tower as high-risk if at least one instance of the condition parameter is at high-risk. In this thesis, MIL using all detected regions for each tower was not possible due to limited GPU memory. Hence, sub-bags of instances from each tower were classified.

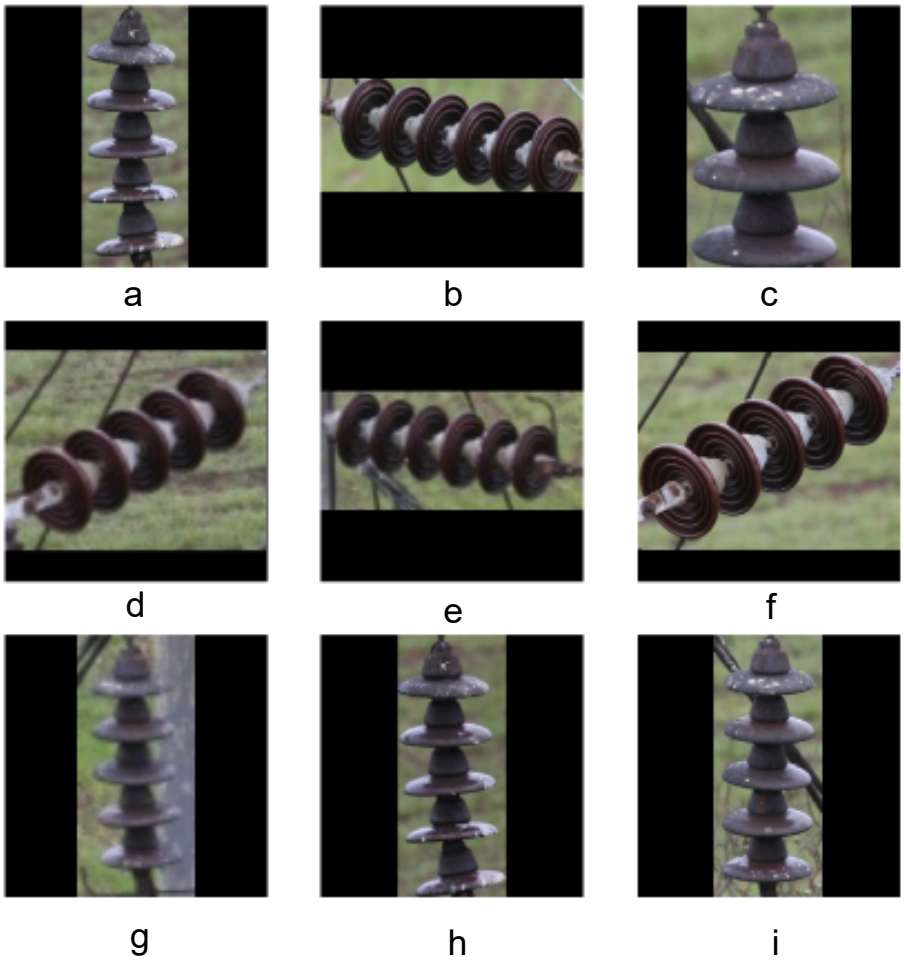


Fig. 6.17 Uneven rusty instances of porcelain insulators supported by the same tower. (b), (d)-(f) are healthy instances.

Normally, the number of tower views (images) should be controlled to limit the number of sub-image regions. With such control in place, there may be no need for aggregating sub-bag class predictions as suggested. A sub-bag of 8 instances was used and each sub-bag remained within either training or testing split to avoid bias. Increasing sub-bag size from 8 to 16 slightly improved AUC from 0.90 to 0.92 on insulator-based tower condition classification.

A MIL could be used for tower condition classification and relying on key instances, e.g., using insulator U-bolts selected from each tower by means of detection probability scores.

Instance-based outperformed MIL-based method on insulators but there is no significant difference between them using insulator U-bolts (see Fig. 6.13). However, at 40% false-positive rate, MIL-50 and instance-based classifiers achieve high sensitivity and clearly outperformed MIL-16 (Fig. 6.13). The results on 3,647 towers (Fig. 6.14) show that towers could be reliably classified based on the conditions of insulator U-bolts with up to 98% certainty.

The results of both instance and MIL based methods are encouraging considering the variability and diversity of the condition parameters. The methods were evaluated using a real-world dataset that was not created for these experiments, suggesting the effectiveness of the automated tower condition classification pipeline. A major contribution of this Chapter is the classification of OHL towers without explicitly labelling all the instances of either insulators or U-bolts they support. The following are some useful design considerations.

1. Towers could be single or double circuit and condition labels were assigned to the circuits separately. In this thesis, a single label derived by taking the maximum condition rating of both circuits was used.

2. Tower level labels were provided. During training of instance classifiers, a tower label was transferred to all the instances detected from it and used for instance classification.

3. MIL classifier needs a bag of instances. It is possible to train a MIL using selected instances from each tower and still achieve good classification results. To utilise all detected instances, MIL of sub-bags was introduced for training on limited GPU memory.

Towers are usually painted to reduce rate of corrosion. Metallic tower structures are particularly noted for this failure mode resulting to components failure. The classification of tower paintwork will be investigated in Chapter 7.

Chapter 7

Images-based Classification of Tower Paintwork Deterioration

7.1 Introduction

Steel towers are painted to protect against direct environmental influence but over time and depending on location, tower paintwork degenerates turning from a silvery-gray colour to red, yellow or red-brown. The breakdown of painting on electrical overhead line towers accelerates components failures such as insulator and U-bolts. Corrosion is further accelerated by high humidity and as such is more prevalent along coastal lines and in forests. Standard industry practice to remedy rusty towers is to apply painting to affected towers. With regular paintings, the useful life of towers can be indefinite [20] but if painting is done too late, additional surface preparation would be required. It is therefore desirable to identify towers with this failure mode to schedule for the next refurbishment window. Current approach involves taking multiple aerial images and classifying the level of paint defects, albeit manually. It is difficult to manually quantify the area covered by paintwork defects from images. The process is costly and subjective.

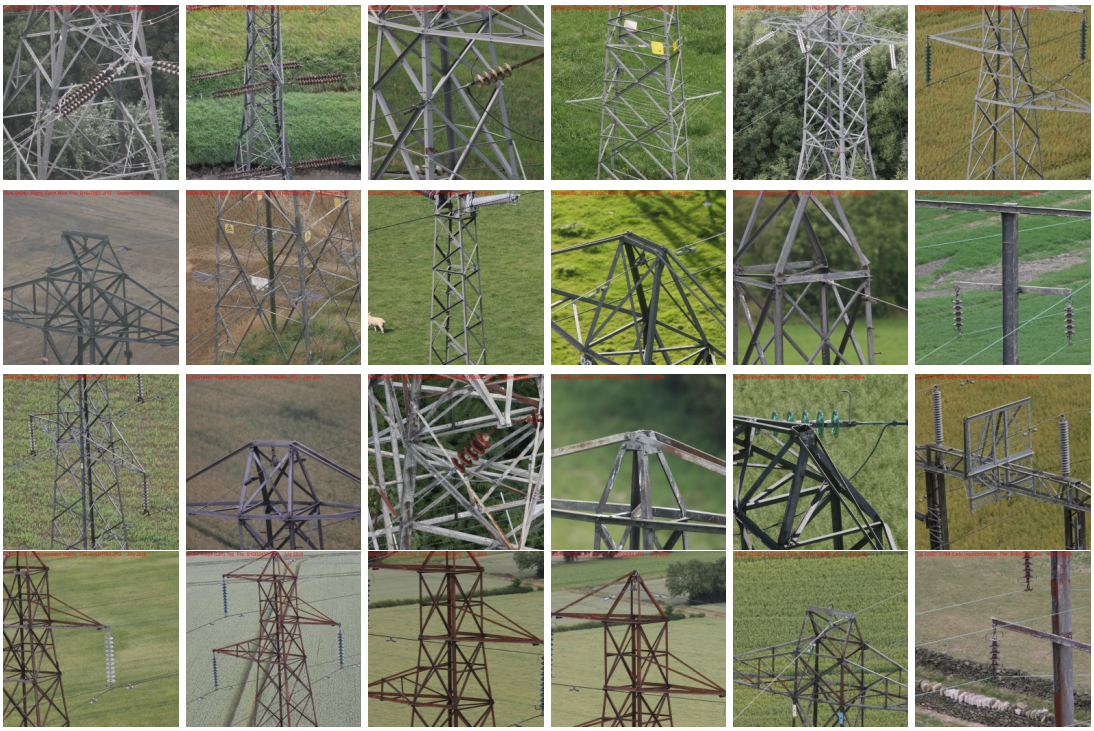


Fig. 7.1 Paintwork deterioration levels. Rows 1, 2, 3 are examples of towers within normal paintwork operating conditions and labelled CR-1, CR-2 and CR-3, respectively. Row 4 are examples of towers at high-risk of paintwork failure - labelled CR-4.

We demonstrated in Chapter 6 that tower conditions can be reliably classified from aerial images and without explicitly labelling instances of insulators and U-bolts. In this Chapter, we consider a different condition parameter involving tower paintwork. We leveraged a real inspection dataset, which is representative tower conditions covering a range of component types, scenes, environmental conditions, and viewpoints. Multiple images of each tower were provided but there is one label (condition rating). We assigned the tower label to all its images and performed supervised learning. Given the high number of images in our dataset, which is representative of the diverse nature of this failure mode and their inspection conditions, we aim to determine if paintwork degradation can be learned directly from aerial images of towers. The main contribution of this paper is the classification of tower paintwork defects. Learning to associate this failure mode from whole tower images would speed up processing and eliminate the need for sub-image annotations.

7.2 Experimental Methods

There are 4 classes of paintwork defects in our dataset. Each class is the condition rating (CR) of an entire tower structure. Images of towers at different CRs are shown in Fig. 7.1. Table 7.1 shows the towers and images distribution. CR-1 are tower with relatively new paintings, CR-2 represents towers with less than 10% paintwork wear, or with primer visible through topcoat, CR-3 is for moderate paint wear (i.e., between 10% and 50% of defects like cracking and peeling). The first three classes are within normal operating conditions and would not pose any risk. CR-4 is a more severe case (high-risk) of paintwork failure with more than 50% wear or paint peeling, AND/OR any sign of paint blistering with evidence of severe rust underneath paint. This high-risk category would require urgent painting. Table 7.1 does not include unpainted towers. Paintwork deterioration can affect both suspension and tension tower types and were used in this experiment. The data set was split into 75% training, 12.5% testing and 12.5% for validation. Considering that multiple images were taken from

Table 7.1 Distribution of towers and images for paintwork classification

Split	CR	Class label	Towers	Images
Training	CR-1	Normal	460	26,903
	CR-2		1,767	101,631
	CR-3		1,500	84,133
	CR-4	High-risk (positive)	1,021	45,027
Validation	CR-1	Normal	77	4,588
	CR-2		295	16,728
	CR-3		250	13,387
	CR-4	High-risk (positive)	170	7,586
Testing	CR-1	Normal	77	4,496
	CR-2		295	16,623
	CR-3		250	14,462
	CR-4	High-risk (positive)	171	7,496
Total			6,333	343,060

each tower, splitting was done at tower level. This ensures that all the images from a tower are within the same split.

The original images in our dataset have dimensions 5616×3744 pixels and tower structures are mostly central. We resized each image to input size of 256×256 pixels, making sure the aspect ratio is maintained. First, we resized along the shorter side of the image (height) and then cropped along the centre. Parts of the image (mostly the background) were discarded during cropping leaving substantial amount of the tower structure at the middle (Section 3.4.1). It is useful practice to add randomness to training data such that the network samples relatively different sets of images at each step and hopefully avoid over-fitting. Sub-images (224×224 pixels) were cropped from any random position on the original inputs (Fig. 3.5). In addition, we applied horizontal flipping and zoomed images by 20%. All input images were normalised using RGB mean subtraction. Mean subtraction involves computing the channel-wise mean of all the cropped object images, i.e., mean of red channel, mean of green channel, and mean of blue channel. The computed mean value was subtracted from the corresponding channel in each image. RGB mean subtraction is a useful way to reduce the effect of varying illumination.

There are 6,333 towers in the dataset for training and evaluation of the method. As shown in Table 7.1, there are 212,667 images from towers that are within normal operating conditions and 45,027 images of the high-risk class in the training set. Essentially, this is a binary classification problem with a threshold at CR-3. All condition rating outside the normal category would be actioned. There is a high class imbalance between the two classes. It is common for under-representation of the positive class in a binary classification problem. One way to train a classifier with such challenging class imbalance might be to generate synthetic data using the original examples to make up for the difference. Tao et al. [115] generated 996 synthetic images from 60 original images. Alternatively, class imbalance

could be checked by penalising the loss of the smaller class more than the dominant class.

$$w_j = n/kn_j \tag{7.1}$$

In this experiment, we used a balanced class weights parameter defined in Eqn. 7.1 [88] during training to penalise the loss of the high-risk class, where w_j is the class weight for class j, n is the total number of samples, n_j is the number of samples of class j and k is the number of classes. The class weights for normal and high-risk classes are 0.6059 and 2.8615, respectively. This implies that the loss of a high-risk class weighs approximately 5 times of the normal class. Eqn. 7.1 may be implemented as a function and called during training. Alternatively, the computed values could be supplied for training as a dictionary of class-index and weight pair.

We fine-tuned EfficientNetB0 for tower paintwork classification as shown in the block diagram (Fig. 7.2). As shown, the base network up to the top convolutional (penultimate) layer was used. The output of the last conv. was flattened, connected to an FC-layer with 128 nodes, followed by a regularisation term of 50% dropout. This new layer effectively reduces the dimensionality of the model before a Softmax output. With input size of 224×224 pixels, the final network has 12M parameters.

Fig. 7.3b shows a modified ResNet [99]. As shown in Fig. 7.3a, there are 4 residual blocks in the standard ResNet50v2. Each residual block has 3 convolutional layers of 1x1, 3x3 and 1x1, respectively. The residual blocks are based on bottle-neck structure defined by the number of filters in each convolutional layer, e.g., 64, 64, 256. By stacking up residual blocks as shown, a deep network with many layers is made (ResNet50v2 has a total of 53 convolutional layers). Instead of the standard ResNet50v2, a custom network with 3 residual blocks was trained from scratch.

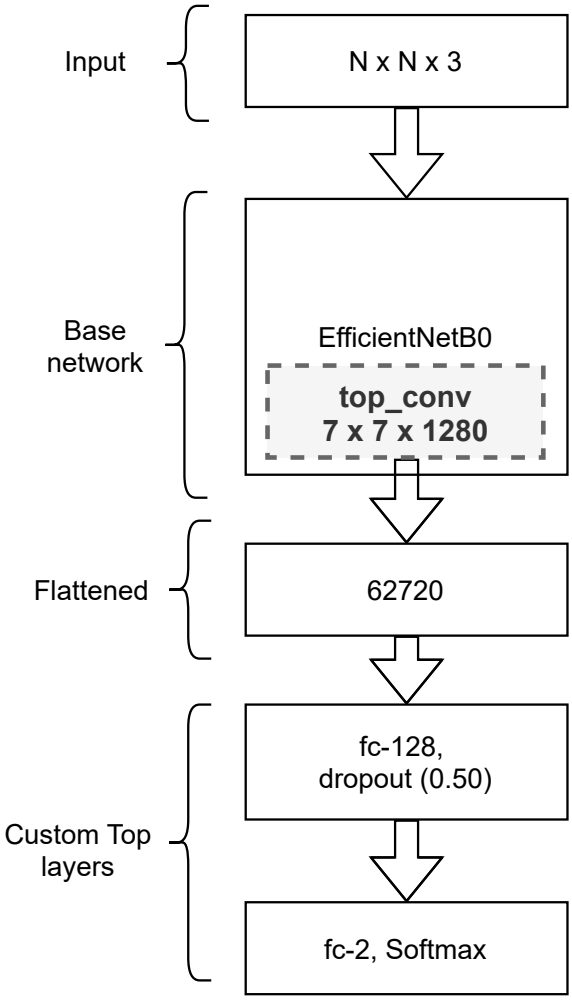


Fig. 7.2 Block diagram of EfficientNet-based model. The top conv. (penultimate) layer was flattened and used as input to a fully connected (FC) layer with 128 nodes and followed by dropout. The output is a Softmax activation.

A binary cross-entropy loss function and stochastic gradient descent with momentum of 0.9 were used. Initial learning rate of 0.001 with a step-decay (Eqn.7.2 [99]) was applied,

$$\alpha = \alpha_i k^{\frac{1+e}{d}} \tag{7.2}$$

where α_i is initial learning rate of 0.001, k is a learning rate decay constant = 0.25, e is the current epoch, and d is the number of epochs before updating the learning rate and is set to 5.

Residual block 1	1x1, 64 3x3, 64 1x1, 256	x 3
Residual block 2	1x1, 128 3x3, 128 1x1, 512	x 4
Residual block 3	1x1, 256 3x3, 256 1x1, 1024	x 6
Residual block 4	1x1, 512 3x3, 512 1x1, 2048	x 3

(a) Building blocks of a standard ResNet50v2

Residual block 1	1x1, 32 3x3, 32 1x1, 128	x 3
Residual block 2	1x1, 64 3x3, 64 1x1, 256	x 4
Residual block 3	1x1, 128 3x3, 128 1x1, 512	x 6

(b) Building blocks of custom ResNet.

Fig. 7.3 Custom ResNet. (a) Standard ResNet50v2 with 4 residual blocks. (b) Custom ResNet with 3 residual blocks.

Table 7.2 shows the network training parameters. As shown, the networks were trained for up to 25 epochs. Training stopped if the validation loss is stagnated for up to 10 epochs. This is a known as early stopping and is a good way to monitor over-fitting. A batch size of 16 images was used. Instead of 10-fold cross-validation, we used a held-out test set comprising of 43k images from 793 towers. We consider this number of towers and images sufficient for evaluating our network performance. Algorithm 4 shows a high-level description of the system.

Table 7.2 Parameter selection and the values used for training image classifiers for paintwork failures. Initial learning rate was scaled according to Eq. 7.2.

Parameter	Value
Number of instances per batch	16
Maximum number of epochs	25
Optimisation function	SGD with momentum of 0.9
Learning rate	0.001
Loss function	Binary cross entropy

Algorithm 4: Image-based tower paintwork classification.

Data: *tower*
Result: *ConditionRating*

$CR \leftarrow 0$;
$N \leftarrow len(tower)$; /* Find the number of images for the tower. */
while $img \in tower$ **do**
  $image \leftarrow imageRead(img)$; /* Read an image of $N \times N$ pixels. */
  $y \leftarrow f(image)$; /* Classify image as normal or highRisk. */
  $pos \leftarrow y(highRisk)$; /* Score of highRisk prediction. */
  $CR := CR + pos$; /* Accumulate all image ratings. */
$ConditionRating \leftarrow \frac{CR}{N}$; /* Returns the aggregate tower rating. */

7.3 Results

High-risk (i.e., positive) towers, CR-4 were classified against towers within "normal" paintwork condition i.e., CR-1, CR-2, and CR-3. At test time, the classifier outputs two probability scores for normal and high-risk classes, respectively. We evaluate by taking the outcomes under the high-risk class. Tower label was then determined by aggregating the classification results of the multiple images of each tower. We used either averaging or the highest score from the set of images from a tower. The aggregate score is considered the final predicted tower paintwork level.

There are 793 towers in the test set (622 towers within normal paintwork condition and 171 at high-risk). A summary of the classifiers performance on the test set is shown in Table 7.3. As shown, averaging predictions always performed better than taking a max for all

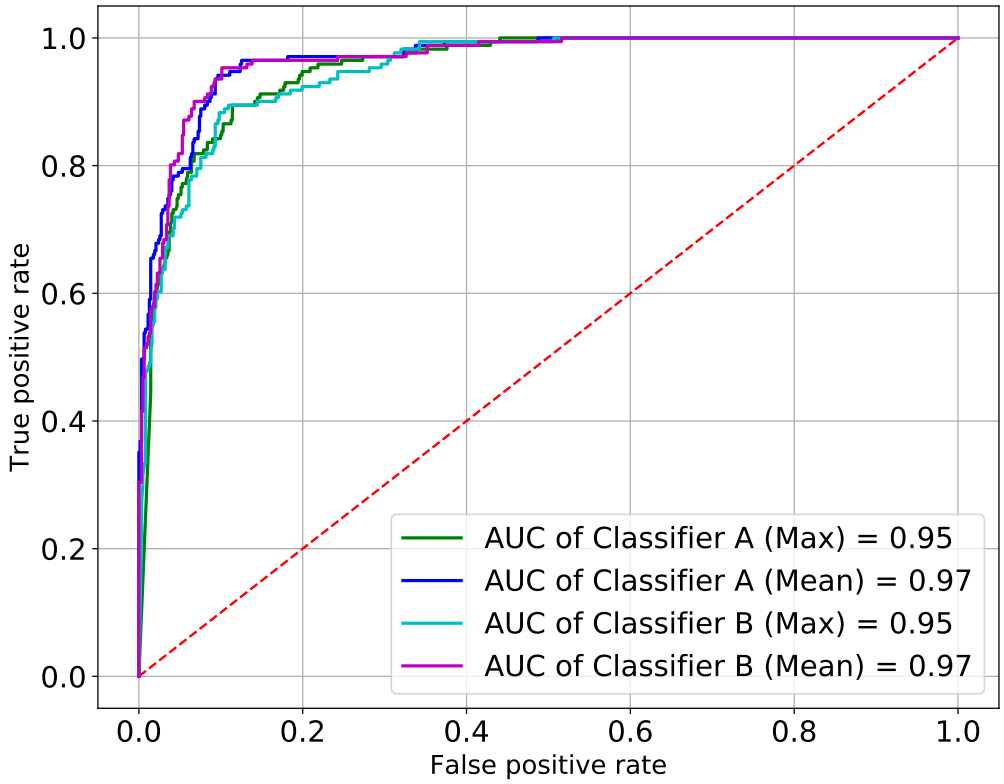


Fig. 7.4 ROC Curves of fine-tuned EfficientNet

the classifiers. As stated in Section 7.2, it is useful practice to add randomness to training data such that the network samples relatively different sets of images at each step and hopefully

Table 7.3 Tower paintwork classification results. A and C were trained using images of 256×256 pixels. B and D used randomly cropped 224×224 pixels of the base input at each training step.

Classifier	Model	Input size	Cropping	Mean	Max
A	EfficientNetB0 + ImageNet	256 x 256		0.97	0.95
B	EfficientNetB0 + ImageNet	224 x 224	☑	0.97	0.95
C	Custom ResNet	256 x 256		0.93	0.90
D	Custom ResNet	224 x 224	☑	0.95	0.92

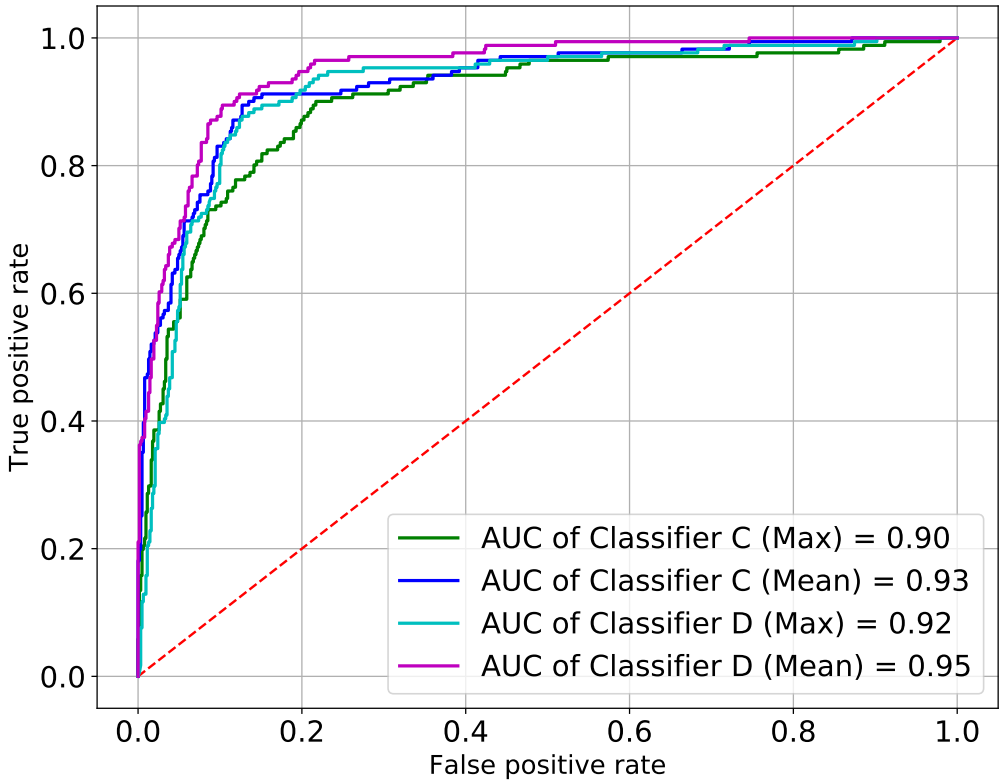


Fig. 7.5 ROC Curves of Custom ResNet

avoid over-fitting. Whereas classifiers A and C were trained using fixed 256×256 pixels, B and D used crop preprocessing with an aim to add jitter to the original input (Table 7.3). Training an EfficientNetB0 with inputs of 256×256 pixels produced the same results with that trained using 224×224 pixels. However, training ResNets with randomly cropped inputs improved the AUC score from 0.93 to 0.95 (classifier D). Fig. 7.4 shows the ROC curves of the classifiers based on fine-tuned EfficientNetB0. As shown in Table 7.4, the best classifiers A & B predicted 91% of the test towers correctly. Inter-classifier reliability test shows 96% agreement at optimal thresholds, showing that similar features were learnt by different networks.

Table 7.4 Confusion matrices on averaging class predictions

Class	Classifier A (Mean)		Classifier B (Mean)	
	Normal	High-risk	Normal	High-risk
Normal	559	63	562	60
High-risk	8	163	11	160

Table 7.5 Expanded confusion matrices

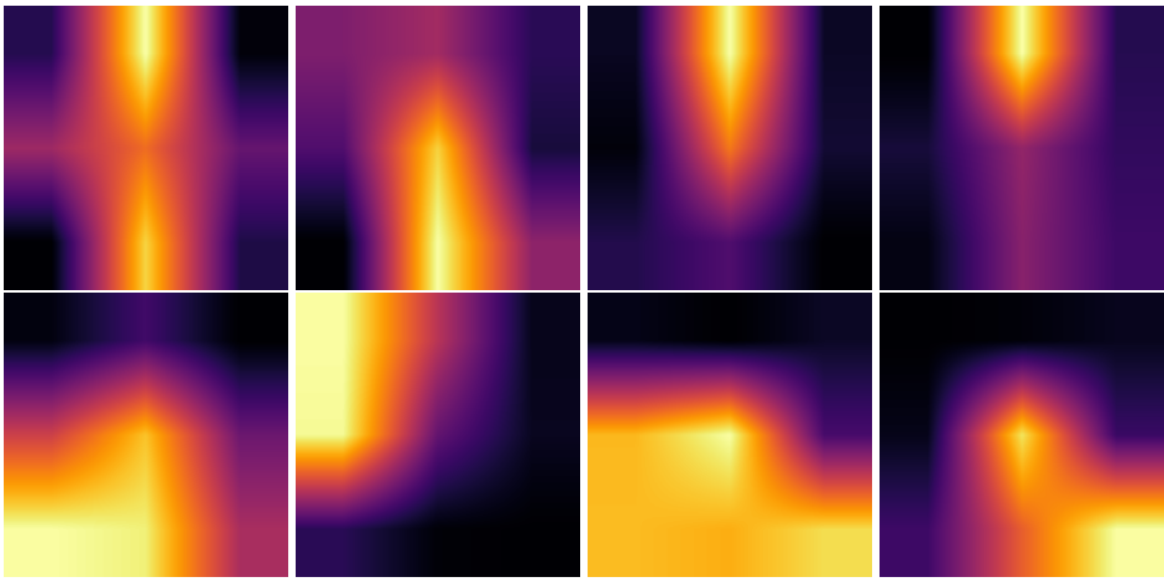
CR	Class	Classifier A		Classifier B	
		Normal	High-risk	Normal	High-risk
1	Normal	76	1	76	1
2		284	11	286	9
3		199	51	200	50
4	High-risk	8	163	11	160

7.4 Summary

Paintwork deterioration can take place anywhere on metallic tower structures irrespective of configuration. This failure mode could be visible on pylons of the tower e.g., the lattice

Fig. 7.6 Images of mis-classified towers. All the images in a row belong to the same tower. Images in Row 1 were labelled CR-1 (Normal class) but mis-classified to be at high-risk. Similarly, images in Rows 2 to 5 were labelled CR-4 (high-risk) but predicted to be in normal operating condition.

(a) Original images



(b) Heat maps

Fig. 7.7 Grad-CAM visualisation. (a) Input images of correctly classified "At-risk" towers, (b) Corresponding heat maps of at-risk towers predicted with high confidence.

structure and cross-arms. As shown in Figs. 7.1, it can be challenging to quantify the level of paintwork degradation on tower images. The aim of this Chapter was to determine if paintwork degradation could be learned directly from whole tower images as it would require substantial investment to classify parts of tower. We classified multiple images without labels

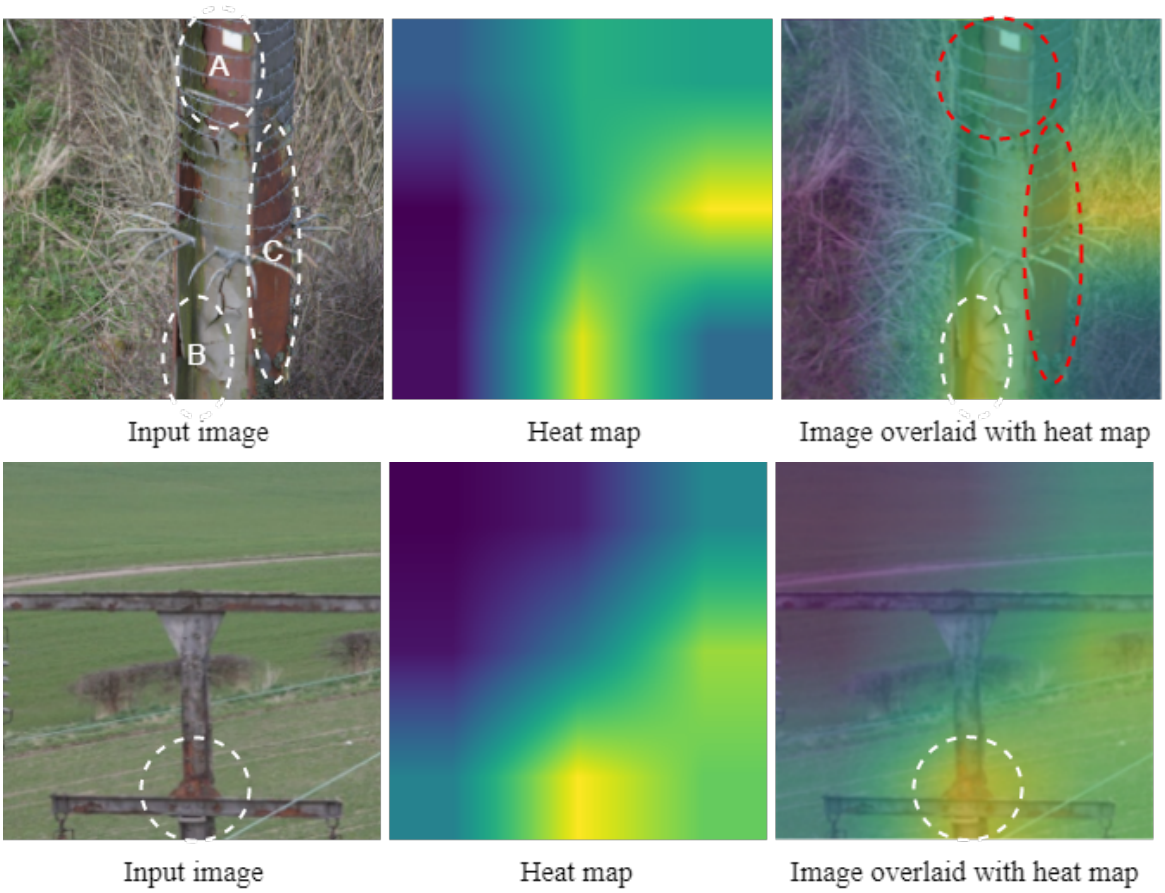


Fig. 7.8 Grad-CAM visualisation (difficult heat maps). LEFT: input images, MIDDLE: corresponding heat maps, RIGHT: overlaid each image with corresponding heat map.

for the local regions. Tower label was transferred to the multiple images taken from it. At test time, tower label was determined by aggregating image class predictions. We classified the high-risk towers i.e., those labelled CR-4 versus the rest of the towers (CR-1, CR-2, CR-3), which are within normal operating conditions. As shown in Table 7.5, the classification results compare reasonably to ground truths. More towers of CR-3 class were classified as high-risk than CR-1 and CR-2 as would be expected.

As shown in Table 7.5, classifiers A and B missed a tower rated CR-1, and this was the same tower in each case. It is encouraging to see that most false-positives occurred for CR-3 towers, with the false positive rate much reduced for towers labelled CR-2 and CR-1. It may be costly to classify normal towers as high-risk but will be more costly to classify

a high-risk tower as normal within an inspection window. Images of some mis-classified towers (especially where the classifiers were in agreement) are shown in Fig. 7.6. The tower illustrated in Row 1 of Fig. 7.6 had 48 images in total; it appears to have been painted recently but was mis-classified as a high-risk tower. Although the pylon is within normal paintwork condition, we observed that the insulators were rusty, and the images were classified at higher probability scores. Averaging multiple image-level predictions resulted in high-risk class. Similarly, all the classifiers were in agreement on eight towers rated CR-4 but mis-classified to be at normal paintwork condition. The images shown in Rows 2 to 5 of Fig. 7.6 do not appear to have up to 50% paintwork defect and may be examples of labelling error. Heavy background clutter is likely to be a contributory factor in some mis-classifications as shown in Row 3 of Fig. 7.6.

Fig. 7.7 shows class activation mappings of the model on some high-risk tower images. As shown, all the heat maps in Fig. 7.7 suggest that paintwork predictions were based on the pylons rather than from background regions (dark corners). Whereas heat maps are a good way to interpret CNN-based image classification, they can be imperfect in some cases and therefore difficult to interpret. Examples of cases where the network appeared to struggle are shown in Fig. 7.8. Row 1 has three regions with defective paintwork marked as A, B, C. The regions marked "A" was not highlighted as much as "B & C". Also, the model highlighted what looks like a reflection of corroded part of "C". In Row 2, image region, which is not part of the pylon was highlighted showing that confusion can come from background.

These results demonstrate the effectiveness of the proposed inspection method for automatic classification of tower paintwork defects. It also shows that tower level labels are adequate for paintwork condition rating.

Chapter 8

Automated Tower Assessment System

8.1 Introduction

Electrical overhead line assets are mostly ageing and require continuous assessment and refurbishment to remain useful. The Condition Based Risk Management model (CBRM) [20] is a popular network asset management model within the industry. The model allows for individual components rating and then enables a collective economic impact analysis for the medium and long term. The main components of CBRM involves aerial survey of the assets resulting to the selection of "At-risk" candidates for additional assessment. Fig. 8.1 shows a block diagram of the assessment and refurbishment process adopted at the Northern Powergrid [20]. The model has two stages of assessment that could inform if a tower requires intervention namely routine and detailed assessments.

Routine assessment is the first step and would normally involve non-intrusive aerial surveys of scheduled electrical network segments. The result of routine inspection sometimes would lead to additional detailed investigation, which could then lead to intervention. The effectiveness of the inspection and refurbishment model would ultimately depend on how quick and accurate the routine assessment is able to highlight areas of the network to channel limited resources.

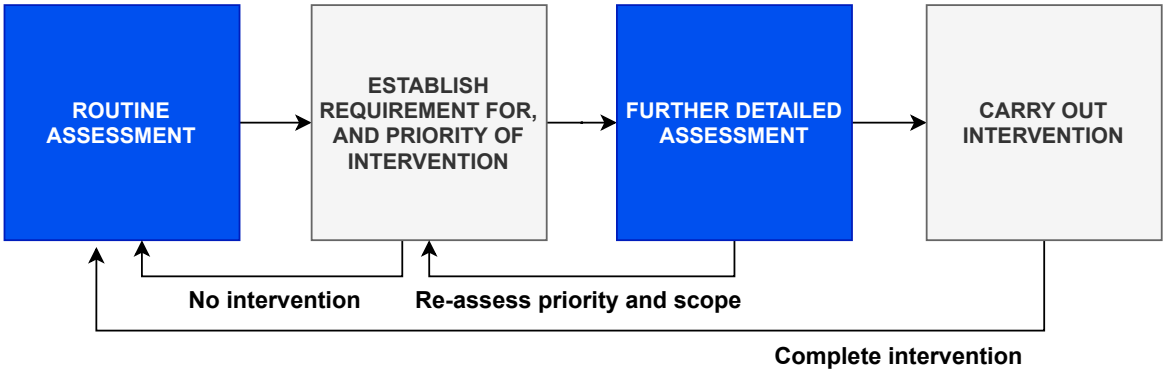


Fig. 8.1 Tower inspection and refurbishment workflow

8.2 Routine Assessment

This thesis presents the automation of tasks within the routine assessment stage of the workflow in Fig. 8.1. Automating the process will facilitate the aerial surveys and reduce the cost of analysing the information from line inspection. There are over fifty inspection parameters that are critical along a network segment and more specifically on individual towers. These parameters include vegetation encroachment, line sagging, corrosion on tower and insulators. We selected three failure modes namely, insulators, U-bolts and paintwork for the evaluation of our proposed method. Four sub-systems were developed namely tower type classification, instance detection, tower condition classification and finally, tower paintwork classification. It is important to highlight that tower labels were provided and not labels for individual components.

8.2.1 Classification of tower types

Tower type detection is a precursor for identifying individual components. The reason this is necessary is because the failure modes investigated may be found on some tower type. For example, to rate a tower on the condition of U-bolts would require identifying suspension towers only. Ability to filter information and process only what is necessary could lead to improvement in response time of the routine inspection process. Classification of tower

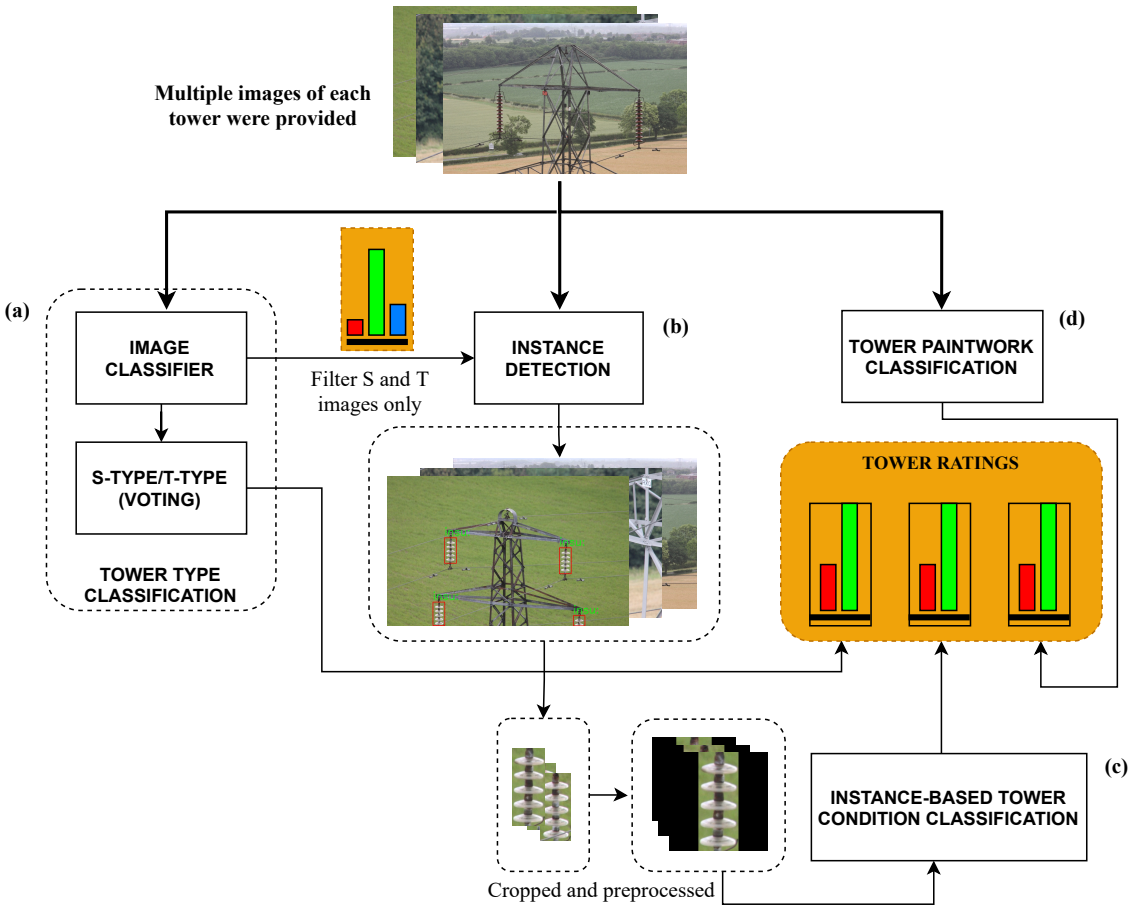


Fig. 8.2 Block diagram of automated routine assessment system. (a) Tower type classification, (b) Detection of instances of insulators and U-bolts, (c) Tower condition classification and (d) Tower paintwork classification.

types was described and evaluated in Chapter 4. It is often the case that components such as U-bolts and insulators are found around the cross-arms instead of body and foot regions. Therefore, a first step to analysing a tower defect based on instances of these components would require filtering images with cross-arms, i.e., S and T images as shown in Fig. 8.2, (a).

8.2.2 Automatic detection of insulators and U-bolts

As shown in Fig. 8.2, (b), instance detection used images identified as S and T. Considering that multiple images of the tower are taken during an aerial survey, it is important to filter images of tower regions that are likely to have the object of interest. If the inspection

parameter is to identify concrete muffs on the foot of towers, it would be unnecessary to compute images of tower peak.

Chapter 5 presented two CNN-based detectors (Mask R-CNN and RetinaNet) for this application and RetinaNet achieved AP of 96.4%, which is comparable to the two-stage model. In addition, RetinaNet achieved a run-time speed that is 3× faster than Mask R-CNN. The utility of Fig. 8.2, (b), was demonstrated in Chapter 6 in which insulator and U-bolt instances were detected from towers images that are disjoint from those used in developing the detection networks.

8.2.3 Tower rating using detected instances of insulators and U-bolts

As shown in Fig. 8.2, an instance detector was used to extract components like insulators in all the images of identified tower types. The detected regions were then cropped, resized and padded with zeros to form a bag of instances and used as inputs to tower condition classification (Fig. 8.2, (c)).

In this thesis, we focus on insulators and U-bolts as exemplar cases of components of interest. As mentioned in Chapter 6, 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 rated CR-4 or CR-5 will be scheduled for immediate detailed intrusive inspection to ascertain the level of intervention.

Two methods for developing tower condition classifiers from bags of component instances were compared. A first approach uses 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. Instance-based classifiers outperforms the MIL of sub-bags techniques described in this thesis.

8.2.4 Images-based classification of tower paintwork deterioration

Fig. 8.2, (d) classifies towers based on the degradation of paintwork. This sub-system accepts whole tower image as input to determine if the tower is at high-risk. The utility of this sub-system was demonstrated in Chapter 7 using over 300k images from 6,333 towers. Aggregating predictions of multiple images of each tower improved classification results.

Chapter 9

Conclusion and Future Work

9.1 Summary of Contributions

Deep learning methods have proved to provide useful tools capable of learning image representation such as presented in this thesis. Assessing OHL tower conditions from aerial images is a challenging task. Considering the vast number of images to be inspected, the proposed methods can provide prognostic information and ultimately facilitate faster and cost-effective results.

The main contributions of this thesis are:

1. **Classification of tower functions as precursor for asset inspection**: Chapter 4 presented the classification of towers as suspension or tension. This could be considered as a prerequisite for tower condition classification. The classification of towers as suspension or tension might be helpful for identifying components on a tower. The study revealed the characteristics of images that led to good classification results. Images that are close-up, clear and with single towers contributed to successful results. In contrast, images with multiple towers, long-range and cluttered background were responsible for mis-classifications.

2. **Detection of insulators and U-bolts**: Chapter 5 investigated automated extraction of OHL assets for condition assessment. This is the first published work on automated detection of U-bolts from aerial images. We found that multi-task detection improved performance and should be used.

3. **Tower condition classification**: Instances of condition parameters on the tower namely insulator mechanical rust and corrosion of insulator U-bolts were investigated in Chapter 6. It demonstrated that failure modes of OHL towers could be reliably classified from aerial images. Tower classification could be achieved without explicitly labelling all the instances of the condition parameter. Importantly, while existing studies classified sub-image regions of towers representing either insulators or U-bolts, we classified a whole tower. Our approach achieved tower condition classification without explicitly labelling instances of the condition parameters. We show that tower level labels are adequate for the task thereby reducing the requirement of fine-grained data annotation.

4. **Single label for dual circuit towers**: Tower condition ratings based on insulators and U-bolts were provided separately for left and right circuits. Left circuit denotes components on the left when facing the tower structure. Components on the right of the tower view belong to the right circuit. However, there are single circuit towers in the industry dataset used in our experiments. For dual circuit towers, we used a single label by picking the highest condition ratings. Combining condition ratings enabled the inspection of both single and double circuit towers simultaneously. This would speed up analysis and ensure a more generic system.

5. **Classification of tower paintwork deterioration**: Previous methods classified localised parts of a tower with paintwork defects. It would require a substantial investment to annotate all parts of a tower with this defect. We classified multiple images

without labels for the local regions and aggregated image class predictions. The results demonstrate the effectiveness of the proposed inspection method and that we do not need to annotate every region on a tower with paintwork failure.

9.2 Limitations

There are some limitations arising from the data set and the assessment pipeline.

1. Tower condition classification considered identifying assets that are at risk against those within normal operating condition to facilitate processing. Condition ratings were merged using CR-3 as threshold point such that CR-4 and CR-5 are positive and CR-1, CR-2 and CR-3 are treated negative class. This will direct efforts and limited resources to additional intrusive or non-intrusive assessment. It will be useful to detect condition ratings separately.

2. High voltage towers are managed on individual circuit basis. The methods discussed in this thesis classified the tower and would require additional information to identify condition parameters on circuit by circuit basis.

3. Classification of sub-bags was a work-around for learning a MIL with limited GPU memory. Tower label was determined by aggregating sub-bags predictions. The amount of memory available determines for example, the batch size used for image classification task. We experimented with inputs of 224×224 pixels and handling bags of all extracted object instances per tower resulted to out-of-memory (OOM) error. Therefore, our MIL was trained using sub-bags of instances.

4. The images used in our experiments were taken using helicopter services, but our method could be deployed from a UAV. These are different aerial platforms with different flight pattern and speed which might affect image quality.

9.3 Future Work

The last two decades have witnessed a high volume of research focusing on the use of image processing techniques including deep learning for applications in many domains and specifically for the automation of OHL assets inspection. Research has demonstrated that deep learning can deliver high precision in image analysis for detection, classification, and segmentation tasks. Often, there is a lack of properly labelled data in the right quantity to satisfy the requirement of DL methods. There is need for a standardised public dataset of OHL inspection to promote industry-wide collaboration and improvement of inspection methods.

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 (UAV). A few hours of helicopter inspection service could buy a fully equipped UAV. In addition to lower cost, they are small and more flexible for aerial photography. The use of UAVs for aerial photography could lead to quick collection of short-range OHL assets data. The advantage of short-range images is that objects, especially smaller components could be reliably identified/detected. However, there might be issues arising from remote piloting the aircraft. UAV autonomy may improve the capacity for sensing overhead lines. Object tracking, collision detection and avoidance are examples of useful features to be considered for the inspection pipeline. DNOs are looking at the demands of the future with digitisation and automation of OHL inspection as major drivers for efficient energy service delivery.

The methods presented in this thesis were based on tower images taken from helicopters. We assume that because of size of the aircraft and safety concerns, photographs would have been taken at reasonable distance from the towers. There could be some performance issues when deployed on a different platform because of the variation in flight patterns and image quality. There are techniques for enhancing image quality, which can be built into the image

analysis pipeline for data correction functions. For example, [12] enhanced blurred images using Super-Resolution Convolutional Neural Network (SRCNN) before detecting insulators from could be a useful improvement for the proposed tower assessment pipeline.

We used Python programming, Tensorflow and Keras deep learning frameworks. Hardware include an NVIDIA RTX2080 GPU with 10GB memory and 4TB storage. Multiple instance learning is computationally demanding. Our deep MIL network has 8M parameters and MILs are generally demanding. All the data elements of each tower should be treated as a batch (bag). In our example, a bag with 50 images of size $224 \times 224 \times 3$ pixels could be accommodated without out-of-memory (OOM) errors. The number of components processed for each tower depends on the number of images sampled from it. Some towers in our dataset were photographed 100 times resulting in the detection of too many instances of insulators and U-bolts. Extracting up to 600 instances per tower may not be required for the task. It would be useful to know how many instances are required to achieve good result. The MIL classifiers might improve if all instances were used simultaneously instead of aggregating sub-bags.

There are several components inspected by distribution network operators on the tower. We proposed methods for automated inspection of towers and evaluated the techniques using several tower failure modes i.e., tower paintwork defects, rusty insulators, and U-bolts. We demonstrated that tower conditions can be efficiently classified and show that deep learning is effective for analysing tower aerial images. Furthermore, the thesis found that fine-grained annotations of towers are not required to achieve good results. Tower labels are adequate for assessing different tower failure modes thereby reducing the cost of the inspection. Identifying towers at risk as presented provides pointers to areas of immediate attention but cannot highlight individual tower condition ratings. Identifying individual classes of tower conditions is important especially for projecting the remaining useful life of the asset.

Therefore, future work will focus on the following:

1. Development of multi-class classifiers to identify individual condition ratings.

2. Development and testing used images acquired using the similar protocols i.e., helicopter service, weather condition and camera settings. It might be useful to demonstrate the usefulness of the integrated tower assessment system using data from a different platform e.g., unmanned aerial vehicles, to check for generalisation of the method.

3. Datasets in this application domain are private and that makes it difficult for researchers to evaluate and compare methods. It will be useful to have a publicly available dataset for this application.

4. Compare more CNNs for the task of instance detection and tower classification, and the effect of different sample sizes on performance.

5. We used EfficientNet as a base network for our MIL classifier. Future work could experiment with other lightweight networks.

6. The first option for solving our MIL problem would be to increase GPU memory capacity from 10GB to say 24GB. This memory capacity might be sufficient to experiment with bags of 256 object instances. Gao et al. [25] found that a batch of 256 images of the same size would need 24GB of GPU space while training a ResNet50, which provides a starting point for further investigation.

7. There are alternative data processing services like Google Cloud AI, Azure, and Amazon AWS ML, which could speed up training of the MIL classifiers. However, because the dataset used in this thesis is the private property of NPG, a written permission will be required to explore this option.

Bibliography

[1] Ahmad, R. and Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers and Industrial Engineering*, 63(1):135–149.

[2] Ahuja, S. K., Shukla, M. K., and Ravulakollu, K. K. (2019). Surface corrosion grade classification using convolution neural network. *International Journal of Recent Technology and Engineering (IJRTE)*.

[3] Aijazi, A. K., Malaterre, L., Tazir, M. L., Trassoudaine, L., and Checchin, P. (2016). Detecting and Analyzing Corrosion Spots on the Hull of Large Marine Vessels Using Colored 3D Lidar Point Clouds. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-3.

[4] Andersson, G., Donalek, P., Farmer, R., Hatziargyriou, N., Kamwa, I., Kundur, P., Martins, N., Paserba, J., Pourbeik, P., Sanchez-Gasca, J., Schulz, R., Stankovic, A., Taylor, C., and Vittal, V. (2005). Causes of the 2003 Major Grid Blackouts in North America and Europe, and Recommended Means to Improve System Dynamic Performance. *IEEE Transactions on Power Systems*, 20(4):1922–1928.

[5] Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., Flynn, D., Elizondo-Gonzalez, S., and Wattam, S. (2020). Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews*, 130:109899.

[6] Atha, D. J. and Jahanshahi, M. (2018). Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Structural Health Monitoring*, 17:1110 – 1128.

[7] Bhujade, R., Adithya, Hrishikesh, S., and Balamurali, P. (2013). Detection of power -lines in complex natural surroundings. In *ICIT 2013*.

[8] Cai, J. and Walker, R. (2010). Height estimation from monocular image sequences using dynamic programming with explicit occlusions. *IET Computer Vision*, 4(3):149.

[9] Carranza-García, M., Torres-Mateo, J., Lara-Benítez, P., and García-Gutiérrez, J. (2020). On the Performance of One-Stage and Two-Stage Object Detectors in Autonomous Vehicles Using Camera Data. *Remote Sensing*, 13(1):89.

[10] Carse, J. and McKenna, S. (2019). Active learning for patch-based digital pathology using convolutional neural networks to reduce annotation costs. In Reyes-Aldasoro,

C. C., Janowczyk, A., Veta, M., Bankhead, P., and Sirinukunwattana, K., editors, *Digital Pathology*, pages 20–27, Cham. Springer International Publishing.

[11] Cerón, A., Mondragón, I., and Prieto, F. (2018). Onboard visual-based navigation system for power line following with UAV. *International Journal of Advanced Robotic Systems*, 15(2):1–12.

[12] Chen, H., He, Z., Shi, B., and Zhong, T. (2019). Research on Recognition Method of Electrical Components Based on YOLO V3. *IEEE Access*, 7:157818–157829.

[13] Chen, M. and Xu, C. (2020). Bird's nest detection method on electricity transmission line tower based on deeply convolutional neural networks. In *4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, volume 1, pages 2309–2312.

[14] Clarke, N. (2011). Asset Monitoring, Management and Optimization. Technical report, Tessella plc.

[15] Couraud, B., Norbu, S., Andoni, M., Robu, V., Gharavi, H., and Flynn, D. (2020). Optimal residential battery scheduling with asset lifespan consideration. In *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 630–634.

[16] Dickie, R., Roman, D., Flynn, D., and Robu, V. (2017). A review of the role of prognostics in predicting the remaining useful life of assets. 27th European Safety and Reliability Conference, ESREL; Conference date: 18-06-2017 Through 22-06-2017.

[17] Du, J., Yan, L., Wang, H., and Huang, Q. (2018). Research on grounding grid corrosion classification method based on convolutional neural network. In *MATEC Web of Conferences*, volume 160.

[18] Dutta, A. and Zisserman, A. (2019). The VIA annotation software for images, audio and video. In *27th ACM Int. Conf. on Multimedia*. ACM.

[19] Dutta, T., Sharma, H., Vellaiappan, A., and Balamuralidhar, P. (2015). Image analysis-based automatic detection of transmission towers using aerial imagery. *Pattern Recognition and Image Analysis*.

[20] Eyre-Walker, R., Howarth, G., Ahmed, R., Lewin, J., and Higinbotham, W. (2015). Application of Advanced Condition Assessment and Asset Management Techniques on Steel Tower Overhead Line Electricity Networks. In *12th International Conference on Transmission and Distribution Construction, Operation and Live-Line Maintenance*.

[21] Fang, S., Mingze, Z., Sheng, L., Xiaoyu, W., and Haiyang, C. (2020). Fast detection method of insulator fault based on image processing technology. In *5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pages 400–406. IEEE.

[22] Fangzheng Zhang, Wanguo Wang, Yabo Zhao, Peng Li, Qiaoyun Lin, and Lingao Jiang (2016). Automatic diagnosis system of transmission line abnormalities and defects based on UAV. In *4th International Conference on Applied Robotics for the Power Industry (CARPI)*, pages 1–5. IEEE.

[23] Fuentes, R., Chapman, T., Cook, M., Scanlan, J., Li, Z., and Richardson, R. C. (2017). Briefing: UK-RAS white paper in robotics and autonomous systems for resilient infrastructure. *Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction*, 170(3):72–79.

[24] Gao, F., Wang, J., Kong, Z., Wu, J., Feng, N., Wang, S., Hu, P., Li, Z., Huang, H., and Li, J. (2017). Recognition of insulator explosion based on deep learning. In *14th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pages 79–82.

[25] Gao, Y., Liu, Y., Zhang, H., Li, Z., Zhu, Y., Lin, H., and Yang, M. (2020). Estimating GPU Memory Consumption of Deep Learning Models. In *28th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 1342—-1352. ACM.

[26] Girshick, R. (2015). Fast R-CNN. In *International Conference on Computer Vision (ICCV)*.

[27] Girshick, R. B., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587.

[28] Golightly, I. and Jones, D. (2005). Visual control of an unmanned aerial vehicle for power line inspection. In *12th International Conference on Advanced Robotics*, pages 288–295. IEEE.

[29] Gowda, S. N. and Yuan, C. (2019). Colornet: Investigating the importance of color spaces for image classification. In Jawahar, C., Li, H., Mori, G., and Schindler, K., editors, *Computer Vision – ACCV*, pages 581–596. Springer International Publishing.

[30] Guo, Q. and Hu, X. (2018). Power line icing monitoring method using binocular stereo vision. In *Proceedings of the 12th IEEE Conference on Industrial Electronics and Applications*, volume 2018-Febru, pages 1905–1908.

[31] Guo, Q., Xiao, J., and Hu, X. (2018). New keypoint matching method using local convolutional features for power transmission line icing monitoring. *Sensors (Switzerland)*, 18(3).

[32] Haes Alhelou, H., Hamedani-Golshan, M. E., Njenda, T. C., and Siano, P. (2019). A survey on power system blackout and cascading events: Research motivations and challenges. *Energies*, 12(4).

[33] Han, J., Yang, Z., Zhang, Q., Chen, C., Li, H., Lai, S., Hu, G., Xu, C., Xu, H., Wang, D., and Chen, R. (2019). A method of insulator faults detection in aerial images for high-voltage transmission lines inspection. *Applied Sciences*, 9(10).

[34] Hao, J., Wulin, H., Jing, C., Xinyu, L., Xiren, M., and Shengbin, Z. (2019). Detection of Bird Nests on Power Line Patrol Using Single Shot Detector. In *Proceedings - Chinese Automation Congress, CAC*, pages 3409–3414.

[35] He, K., Gkioxari, G., Dollar, P., and Girshick, R. (2017). Mask R-CNN. In *International Conference on Computer Vision (ICCV)*, pages 2980–2988, Los Alamitos, CA, USA. IEEE Computer Society.

[36] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Identity mappings in deep residual networks. In *Computer Vision – ECCV*, pages 630–645, Cham. Springer International Publishing.

[37] Herwitz, S., Johnson, L., Dunagan, S., Higgins, R., Sullivan, D., Zheng, J., Lobitz, B., Leung, J., Gallmeyer, B., Aoyagi, M., Slye, R., and Brass, J. (2004). Imaging from an unmanned aerial vehicle: agricultural surveillance and decision support. *Computers and Electronics in Agriculture*, 44(1):49–61.

[38] Hoang, N.-D. and Tran, V.-D. (2019). Image processing-based detection of pipe corrosion using texture analysis and metaheuristic-optimized machine learning approach. In *Comp. Int. and Neurosc.*

[39] Hong-Bin, Z., Long, H., and Yun-Feng, L. (2020). Target Tracking Method of Transmission Line Insulator Based on Multi Feature Fusion and Adaptive Scale Filter. In *5th Asia Conference on Power and Electrical Engineering (ACPEE)*, pages 1626–1630. IEEE.

[40] Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., and Murphy, K. (2017a). Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3296–3297. IEEE.

[41] Huang, X., Zhang, F., Li, H., and Liu, X. (2017b). An online technology for measuring icing shape on conductor based on vision and force sensors. *IEEE Transactions on Instrumentation and Measurement*, 66(12):3180–3189.

[42] Huang, X., Zhang, H., and Zhang, Y. (2017c). Automatic identification and location technology of glass insulator self-shattering. *Journal of Electronic Imaging*, 26(6).

[43] Huynh, N., Robu, V., Flynn, D., Rowland, S., and Coapes, G. (2017). Design and demonstration of a wireless sensor network platform for substation asset management. *CIRED - Open Access Proceedings Journal*, 2017:105–108.

[44] Ilse, M., Tomczak, J. M., and Welling, M. (2018). Attention-based deep multiple instance learning. *CoRR*, abs/1802.04712.

[45] Ippolito, C., Krishnakumar, K., and Hening, S. (2016). Preliminary results of powerline reconstruction from airborne LiDAR for safe autonomous low-altitude urban operations of small UAS. In *Sensors*, pages 1–3. IEEE.

[46] Jabid, T. and Ahsan, T. (2018). Insulator Detection and Defect Classification using Rotation Invariant Local Directional Pattern. *International Journal of Advanced Computer Science and Applications*, 9(2):265–272.

[47] Jalil, B., Leone, G. R., Martinelli, M., Moroni, D., Pascali, M. A., and Berton, A. (2019). Fault detection in power equipment via an unmanned aerial system using multi modal data. *Sensors*, 19:3014.

[48] Jiang, A., Yan, N., Shen, B., Gu, C., Zhu, H., and Huang, H. (2020). Research on Infrared Image Recognition Method of Power Equipment Based on Deep Learning. In *International Conference on High Voltage Engineering and Application*, pages 1–4. IEEE.

[49] Jones, D. and Earp, G. (2001). Camera sightline pointing requirements for aerial inspection of overhead power lines. *Electric Power Systems Research*, 57(2):73–82.

[50] Jones, D., Golightly, I., Roberts, J., and Usher, K. (2006). Modeling and control of a robotic power line inspection vehicle. In *Conference on Computer Aided Control System Design, IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control*, pages 632–637. IEEE.

[51] Junaid, A., Aamir, S. M., Likun, X., and Nadia, A. (2013). Vegetation encroachment monitoring for transmission lines right-of-ways: A survey. *Electric Power Systems Research*, 95:339–352.

[52] K. Everhart and G. Molnar (2021). Severe power cuts in Texas highlight energy security risks related to extreme weather events – Analysis - IEA.

[53] Kaiming, H., Xiangyu, Z., Shaoqing, R., and Jian, S. (2015). Deep residual learning for image recognition. *CoRR*, abs/1512.03385.

[54] Kandel, I. and Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, 6(4):312–315.

[55] Li, J., Shao, Q., Xue, K., Wang, C., and Hu, W. (2017). The icing-thickness detection of high-voltage transmission line based on machine vision. In *International Conference on Information and Automation, ICIA 2017*, pages 381–385.

[56] Li, X., Su, H., and Liu, G. (2020). Insulator defect recognition based on global detection and local segmentation. *IEEE Access*, 8:59934–59946.

[57] Li, Z., Liu, Y., Hayward, R., Zhang, J., and Cai, J. (2008). Knowledge-based power line detection for UAV surveillance and inspection systems. In *Proc. 23rd Int. Conf. Image and Vision Computing New Zealand*, pages 1–6.

[58] Li, Z., Liu, Y., Walker, R., Hayward, R., and Zhang, J. (2010). Towards automatic power line detection for a UAV surveillance system using pulse coupled neural filter and an improved Hough transform. *Machine Vision and Applications*, 21(5):677–686.

[59] Lin, T., Goyal, P., Girshick, R., He, K., and Dollár, P. (2020). Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):318–327.

[60] Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., and Dollár, P. (2015). Microsoft COCO: Common objects in context.

[61] Ling, Z., Zhang, D., Qiu, R. C., Jin, Z., Zhang, Y., He, X., and Liu, H. (2019). An accurate and real-time method of self-blast glass insulator location based on Faster R-CNN and U-net with aerial images. *CSEE Journal of Power and Energy Systems*, 5(4):474–482.

[62] Liu, C., Liu, Y., Wu, H., and Dong, R. (2015). A safe flight approach of the UAV in the electrical line inspection. *International Journal of Emerging Electric Power Systems*, 16:21.

[63] Liu, L., Tan, E., Yin, X. J., Zhen, Y., and Cai, Z. Q. (2019). Deep learning for coating condition assessment with active perception. In *Proc. of the 3rd High Performance Computing and Cluster Technologies Conference*. ACM.

[64] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C. (2016). SSD: Single shot multibox detector. In *Computer Vision – ECCV*, pages 21–37, Cham. Springer International Publishing.

[65] Liu, X., Miao, X., Jiang, H., and Chen, J. (2020). Review of data analysis in vision inspection of power lines with an in-depth discussion of deep learning technology. *ArXiv*, abs/2003.09802.

[66] Liu, Y., Li, J., Xu, W., and Liu, M. (2017). A method on recognizing transmission line structure based on multi-level perception. *Image and Graphics*.

[67] Liu, Y. and Mejias, L. (2012). Real-time power line extraction from Unmanned Aerial System video images. In *2nd International Conference on Applied Robotics for the Power Industry (CARPI)*, pages 52–57. IEEE.

[68] Maeda, K., Takahashi, S., Ogawa, T., and Haseyama, M. (2017). Automatic estimation of deterioration level on transmission towers via deep extreme learning machine based on local receptive field. In *International Conference on Image Processing (ICIP)*, pages 2379–2383.

[69] Maeda, K., Takahashi, S., Ogawa, T., and Haseyama, M. (2018). Estimation of deterioration levels of transmission towers via deep learning maximizing canonical correlation between heterogeneous features. *IEEE Journal of Selected Topics in Signal Processing*, 12(4):633–644.

[70] Manivannan, S. ., Cobb, C. ., Burgess, S. ., and Trucco, E. (2017). Sub-Category Classifiers for Multiple-Instance Learning and its Application to Retinal Nerve Fiber Layer Visibility ClassificationCategory Classifiers for Multiple-Instance Learning and its Application to Retinal Nerve Fiber Layer Visibility Classification Sub-Category Classifiers for Multiple-Instance Learning and its Application to Retinal Nerve Fiber Layer Visibility Classification. *IEEE Transactions on Medical Imaging*, 36(5):1.

[71] Martinez, C., Sampedro, C., Chauhan, A., and Campoy, P. (2014). Towards autonomous detection and tracking of electric towers for aerial power line inspection. *International Conference on Unmanned Aircraft Systems, ICUAS - Conference Proceedings*, pages 284–295.

[72] Martinez-Sandoval, R., Garcia-Sanchez, A.-J., Garcia-Sanchez, F., Garcia-Haro, J., and Flynn, D. (2014). A comprehensive WSN-based approach to efficiently manage a smart grid. *Sensors*, 14(10):18748–18783.

[73] Matikainen, L., Lehtomäki, M., Ahokas, E., Hyyppä, J., Karjalainen, M., Jaakkola, A., Kukko, A., and Heinonen, T. (2016). Remote sensing methods for power line corridor surveys. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119:10–31.

[74] Miao, X., Liu, X., Chen, J., Zhuang, S., Fan, J., and Jiang, H. (2019). Insulator detection in aerial images for transmission line inspection using single shot multibox detector. *IEEE Access*, 7:9945–9956.

[75] Miguelañez-Martin, E. and Flynn, D. (2015). Embedded intelligence supporting predictive asset management in the energy sector. In *Asset Management Conference*, pages 1–7.

[76] Mokhtar, M., Robu, V., Flynn, D., Higgins, C., Whyte, J., and Fulton, F. (2018). Automated verification of LV network topologies. In *IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6.

[77] Mokhtar, M., Robu, V., Flynn, D., Higgins, C., Whyte, J., Loughran, C., and Fulton, F. (2019). Predicting the voltage distribution for low voltage networks using deep learning. In *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 1–5.

[78] Murthy, V., Tarakanath, K., Mohanta, D., and Gupta, S. (2010). Insulator condition analysis for overhead distribution lines using combined wavelet support vector machine (SVM). *IEEE Transactions on Dielectrics and Electrical Insulation*, 17(1):89–99.

[79] Nguyen, V. N., Jenssen, R., and Roverso, D. (2018). Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *Int. Journal of Electrical Power & Energy Systems*, 99:107–120.

[80] Oberweger, M., Wendel, A., and Bischof, H. (2014). Visual Recognition and Fault Detection for Power Line Insulators. *19th Computer Vision Winter Workshop*.

[81] Odo, A., McKenna, S., Flynn, D., and Vorstius, J. (2020). Towards the automatic visual monitoring of electricity pylons from aerial images. In *15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pages 566–573. Scitpress.

[82] Odo, A., McKenna, S., Flynn, D., and Vorstius, J. B. (2021). Aerial image analysis using deep learning for electrical overhead line network asset management. *IEEE Access*, 9:146281–146295.

[83] Ofgem (2018). RIIO-ED1 Network Performance Summary. Technical report.

[84] Ohta, H., Sato, Y., Mori, T., Takaya, K., and Kroumov, V. (2019). Image acquisition of power line transmission towers using UAV and deep learning technique for insulators localization and recognition. In *23rd International Conference on System Theory, Control and Computing (ICSTCC)*, pages 125–130.

[85] Ortiz, A., Bonnin-Pascual, F., Garcia-Fidalgo, E., and Company, J. P. (2016). Visual inspection of vessels by means of a micro-aerial vehicle: An artificial neural network approach for corrosion detection. In *Second Iberian Robotics Conference. Advances in Intelligent Systems and Computing*. Springer.

[86] Osborne, M., Lantair, J., Shafiq, Z., Zhao, X., Robu, V., Flynn, D., and Perry, J. (2019). UAS operators safety and reliability survey: Emerging technologies towards the certification of autonomous uas. In *4th International Conference on System Reliability and Safety (ICSRS)*, pages 203–212.

[87] Padilla, R., Passos, W. L., Dias, T. L. B., Netto, S. L., and da Silva, E. A. B. (2021). A comparative analysis of object detection metrics with a companion open-source toolkit. *Electronics*, 10(3).

[88] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

[89] Petricca, L., Moss, T., Figueroa, G., and Broen, S. (2016). Corrosion detection using A. I.: A comparison of standard computer vision techniques and deep learning model. In *3rd Int. Conf. on Artificial Intelligence and Applications (AIAP)" Vienna (Austria)*, pages 91–99.

[90] Potnuru, S. P. and Bhima, P. R. (2017). Image processing and machine learning applied for condition monitoring of 11-kV power distribution line insulators using curvelet and LTP features. In *International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, pages 3012–3017. IEEE.

[91] Preeti, C. and Mohan, K. J. (2013). Analysis of Transmission Towers with Different Configurations. *Jordan Journal of Civil Engineering, Volume 7, No. 4*, 7(4):450–460.

[92] Pullum, L., Jindal, A., Roopaei, M., Diggewadi, A., Andoni, M., Zobaa, A., Alam, A., Bani-Ahmed, A., Ngo, Y., Vyas, S., Kumar, R., Robu, V., Flynn, D., Caputo, P., and Rajski Parashis, A. (2017). Big data analytics in the smart grid: Big data analytics, machine learning and artificial intelligence in the smart grid: Introduction, benefits, challenges and issues. Workingpaper, IEEE, United States. IEEE Smart Grid White Paper.

[93] Qiu, Z., Zhu, X., Shi, D., and Kuang, Y. (2020). Recognition of Transmission Line Related Bird Species Based on Image Feature Extraction and Support Vector Machine. In *7th IEEE International Conference on High Voltage Engineering and Application*.

[94] R. Gonzales (2019). PG&E Transmission Lines Caused California's Deadliest Wildfire, State Officials Say.

[95] Redmon, J. and Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6517–6525. IEEE.

[96] Ren, S., He, K., Girshick, R., and Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149.

[97] Robu, V. and Flynn, D. (2017). Outsmart supply dips in renewable energy. *Nature*, page 161.

[98] Rong, S. and He, L. (2020). A joint Faster RCNN and stereovision algorithm for vegetation encroachment detection in power line corridors. In *IEEE Power Energy Society General Meeting (PESGM)*, pages 1–5.

[99] Rosebrock, A. (2018). *Competing in Kaggle: Dogs vs. Cats*, volume 2, pages 105–107. PyImageSearch.

[100] Rosebrock, A. (2019). *Deep Learning for Computer Vision with Python*. PyImageSearch.com, 2.1.0 edition.

[101] Sa, I., Hrabar, S., and Corke, P. (2015). Inspection of pole-like structures using a visual-inertial aided VTOL platform with shared autonomy. *Sensors*, 15(9):22003–22048.

[102] Sampedro, C., Martinez, C., Chauhan, A., and Campoy, P. (2014). A supervised approach to electric tower detection and classification for power line inspection. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1970–1977. IEEE.

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

[104] Sanjay, K. A. and Manoj, K. S. (2018). A Survey of Computer Vision Based Corrosion Detection Approaches. *Technology for Intelligent Systems*.

[105] Santos, T., Moreira, M., Almeida, J., Dias, A., Martins, A., Dinis, J., Formiga, J., and Silva, E. (2017). PLineD: Vision-based power lines detection for Unmanned Aerial Vehicles. In *International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 253–259. IEEE.

[106] Schmidthaler, M. and Reichl, J. (2016). 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. *Computer Science - Research and Development*, 31(3):157–161.

[107] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2019). Grad-cam: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, 128(2):336–359.

[108] Shafiee, M., Zhou, Z., Mei, L., Dinmohammadi, F., Karama, J., and Flynn, D. (2021). Unmanned aerial drones for inspection of offshore wind turbines: A mission-critical failure analysis. *Robotics*, 10(1).

[109] Sharma, H., Bhujade, R., Adithya, V., and Balamuralidhar, P. (2014). Vision-based detection of power distribution lines in complex remote surroundings. In *Proc. Twentieth National Conf. Communications (NCC)*, pages 1–6.

[110] Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1):1–48.

[111] Siddiqui, Z. A., Park, U., Lee, S. W., Jung, N. J., Choi, M., Lim, C., and Seo, J. H. (2018). Robust powerline equipment inspection system based on a convolutional neural network. *Sensors (Switzerland)*, 18.

[112] Simonyan, K. and Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. In *ICLR*, pages 1–14.

[113] Smits, N. (2010). A note on Youden's J and its cost ratio. *BMC Medical Research Methodology*, 10:89 – 89.

[114] Tan, M. and Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 6105–6114. PMLR.

[115] Tao, X., Zhang, D., Wang, Z., Liu, X., Zhang, H., and Xu, D. (2020). Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. *IEEE Trans on Systems, Man, and Cybernetics: Systems*, 50(4):1486–1498.

[116] Thompson, A., Kazemtabrizi, B., Crabtree, C. J., Dao, C., Dinmohamadi, F., and Flynn, D. (2019). Reliability and economic evaluation of high voltage direct current interconnectors for large-scale renewable energy integration and transmission. In *15th IET International Conference on AC and DC Power Transmission*, pages 1–6.

[117] Tian, F., Wang, Y., and Zhu, L. (2015). Power line recognition and tracking method for UAVs inspection. In *Proc. IEEE Int. Conf. Information and Automation*, pages 2136–2141.

[118] Torresani, L. (2014). *Weakly Supervised Learning*, pages 883–885. Springer US, Boston, MA.

[119] Tsutsumi, F., Murata, H., Onoda, T., Oguri, O., and Tanaka, H. (2009). Automatic corrosion estimation using galvanized steel images on power transmission towers. In *Transmission & Distribution Conference & Exposition: Asia and Pacific*, pages 1–4. IEEE.

[120] UK (2002). The Electricity Safety, Quality and Continuity Regulations 2002.

[121] Wang, J., Wang, K., Liu, G., Zhou, W., and Zhou, Z. (2020). Recognition of Defects in Pins Based on Generative Adversarial Network and RetinaNet. *Huanan Ligong Daxue Xuebao/Journal of South China University of Technology (Natural Science)*, 48(2):1–8.

[122] Wang, X., Yan, Y., Tang, P., Bai, X., and Liu, W. (2018). Revisiting multiple instance neural networks. *Pattern Recognition*, 74:15–24.

[123] Xin, L., Jin, H., Tu, Y., Yuan, Z., Lv, Z., and Wang, C. (2020). Defect Detection and Characterization of RTV Silicone Rubber Coating on Insulator Based on Visible Spectrum Image. *IEEE Transactions on Power Delivery*, 35(6):2734–2736.

[124] Yang, T. W., Yin, H., Ruan, Q. Q., Han, J. D., Qi, J. T., Yong, Q., Wang, Z. T., and Sun, Z. Q. (2012). Overhead power line detection from UAV video images. In *Proc. 19th Int. Conf. Mechatronics and Machine Vision in Practice (M2VIP)*, pages 74–79.

[125] Yang, Y., Wang, L., Wang, Y., and Mei, X. (2019). Insulator self-shattering detection: a deep convolutional neural network approach. *Multimedia Tools and Applications*, 78(8):10097–10112.

[126] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? *ArXiv*, abs/1411.1792.

[127] Yu, Y., Cao, H., Wang, Z., Li, Y., Li, K., and Xie, S. (2019). Texture-and-shape based active contour model for insulator segmentation. *IEEE Access*, 7:78706–78714.

[128] Yue, X., Wang, H., and Jiang, Y. (2017). A novel 110 kV power line inspection robot and its climbing ability analysis. *International Journal of Advanced Robotic Systems*, 14(3):1729881417710461.

[129] Zaki, O., Dunnigan, M., Robu, V., and Flynn, D. (2021). Reliability and safety of autonomous systems based on semantic modelling for self-certification. *Robotics*, 10(1).

[130] Zhai, Y., Wang, D., Zhang, M., Wang, J., and Guo, F. (2017a). Fault detection of insulator based on saliency and adaptive morphology. *Multimedia Tools and Applications*, 76(9):12051.

[131] Zhai, Y., Wang, G., Yu, H., and Wei, G. (2017b). Research on the application of the edge detection method for the UAVs icing monitoring of transmission lines. In *International Conference on Mechatronics and Automation (ICMA)*, pages 1110–1114. IEEE.

[132] Zhang, F., Goubran, R., and Straznicky, P. (2012). Obstacle detection for low flying UAS using monocular camera. In *International Instrumentation and Measurement Technology Conference Proceedings*, pages 2133–2137. IEEE.

[133] Zhang, K. and Yang, L. (2019). Insulator segmentation algorithm based on k-means. In *Chinese Automation Congress (CAC)*, pages 4747–4751. IEEE.

[134] Zhang, T., Liu, L., Wiliem, A., Connor, S., Ilich, Z., Van Der Draai, E., and Lovell, B. (2019). Deep corrosion assessment for electrical transmission towers. In *Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–6.

[135] Zhang, Y., Yuan, X., Fang, Y., and Chen, S. (2017a). UAV Low Altitude Photogrammetry for Power Line Inspection. *ISPRS International Journal of Geo-Information*, 6(1):14.

[136] Zhang, Y., Yuan, X., Li, W., and Chen, S. (2017b). Automatic power line inspection using UAV images. *Remote Sensing*, 9:824.

[137] Zhao, X., Robu, V., Flynn, D., Dinmohammadi, F., Fisher, M., and Webster, M. (2019). Probabilistic model checking of robots deployed in extreme environments. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):8066–8074.

[138] Zhou, Z., Yuan, G., Feng, W., Gu, S., and Fan, P. (2020). Target recognition and evaluation of typical transmission line equipment based on deep learning. In Xue, Y., Zheng, Y., and Rahman, S., editors, *Proceedings of PURPLE MOUNTAIN FORUM, International Forum on Smart Grid Protection and Control*, pages 701–709, Singapore. Springer, Singapore.

[139] Zhu, L., Cao, W., Han, J., and Du, Y. (2013). A double-side filter based power line recognition method for UAV vision system. In *International Conference on Robotics and Biomimetics (ROBIO)*, pages 2655–2660. IEEE.

Appendix A

Assessment Report - T7799

T7799-A136			LEFT CIRCUIT				RIGHT CIRCUIT				Hidden free text to fit to page			
Tower Details			nsulators And Fittings				nsulators And Fittings				Defect Image Hyperlinks			
											1	2	3	4
Tower No	**Tension or Suspension (T/S)**	Paintwork	Insulators - Electrical (132kV)	Insulators - Electrical (33kV)	Insulators - Mechanical (Rust)	U bolts or Tower attachments	Insulators - Electrical (132kV)	Insulators - Electrical (33kV)	Insulators - Mechanical (Rust)	U bolts or Tower attachments				
A136(002)	S	3	N	1	3	2	N	N	N	N				
A136(003)	T	4	N	1	2	2	N	N	N	N	B15E1582			
A136(004)	S	4	N	1	4	3	N	N	N	N	B15E1565	B15E1563	B15E1567	
A136(005)	S	4	N	1	3	5	N	N	N	N	B15E1523	B15E1520	B15E1532	B15E1538
A136(006)	S	4	N	1	5	5	N	N	N	N	B15E1492	B15E1517	B15E1504	B15E1514
A136(007)	T	4	N	3	2	2	N	N	N	N	B15E1677	B15E1453	B15E1468	
A136(008)	S	4	N	1	5	4	N	N	N	N	B15E1427	B15E1451	B15E1437	B15E1441
A136(009)	S	3	N	1	5	5	N	N	N	N	B15E1396	B15E1420	B15E1409	B15E1411
A136(010)	T	4	N	1	1	2	N	N	N	N	B15E1369			
A136(011)	S	3	N	1	5	3	N	N	N	N	B15E1330	B15E1338	B15E1342	
A136(012)	T	3	N	1	1	1	N	N	N	N	B15E1707			
A136(013)	T	4	N	1	1	1	N	N	N	N	B15E1279	B15E1716	B15E1268	
A136(014)	S	4	N	1	5	3	N	N	N	N	B15E1723	B15E1242	B15E1244	B15E1252
A136(015)	T	4	N	1	2	2	N	N	N	N	B15E1190			
A136(016)	S	4	N	1	4	5	N	N	N	N	B15E1169	B15E1183	B15E1177	B15E1181
A136(020)	S	4	N	1	2	3	N	N	N	N	B15E1154	B15E1155		
A136(021)	S	4	N	1	2	3	N	N	N	N	B15E1108	B15E1105		
A136(022)	S	4	N	1	2	2	N	N	N	N	B15E1084			
A136(023)	T	4	N	1	2	3	N	N	N	N	B15E1039	B15E1053		
A136(024)	S	4	N	1	5	4	N	N	N	N	B15E1027	B15E1030	B15E1019	B15E1026
A136(025)	S	4	N	1	4	3	N	N	N	N	B15E0986	B15E0999	B15E0994	
A136(026)	S	4	N	1	5	3	N	N	N	N	B15E0956	B15E0971	B15E0960	B15E0967
A136(027)	S	4	N	1	3	3	N	N	N	N	B15E0922	B15E0924	B15E0942	
A136(028)	S	4	N	1	5	3	N	N	N	N	B15E0896	B15E0900	B15E0910	
A136(029)	T	4	N	1	3	3	N	N	N	N	B15E0852	B15E0859		
A136(030)	S	4	N	1	3	3	N	N	N	N	B15E0817	B15E0822		
A136(031)	S	4	N	1	5	3	N	N	N	N	B15E0793	B15E0797		
A136(032)	S	4	N	1	4	3	N	N	N	N	B15E0760	B15E0762	B15E0772	B15E0774
A136(033)	S	4	N	1	5	4	N	N	N	N	B15E0737	B15E0752	B15E0740	B15E0745
A136(034)	T	4	N	1	2	2	N	N	N	N	B15E1825	B15E0708		
A136(035)	S	3	N	1	5	3	N	N	N	N	B15E0674			
A136(036)	S	4	N	1	5	3	N	N	N	N	B15E0641	B15E0649	B15E0653	
A136(037)	S	4	N	1	5	5	N	N	N	N	B15E0604	B15E0616	B15E0619	B15E0623
A136(038)	T	3	N	1	2	2	N	N	N	N				
A136(039)	S	4	N	1	5	3	N	N	N	N	B15E0539	B15E0545	B15E0548	B15E0554
A136(040)	S	3	N	1	5	3	N	N	N	N	B15E0505	B15E0516	B15E0522	
A136(041)	T	3	N	1	2	3	N	N	N	N				
A136(042)	S	3	N	1	5	4	N	N	N	N	B15E0445	B15E1873		
A136(043)	S	3	N	1	3	4	N	N	N	N	B15E0405	B15E0416	B15E0419	B15E0423
A136(044)	S	4	N	1	5	4	N	N	N	N	B15E1884	B15E0386	B15E0391	B15E0394
A136(048)	S	4	N	1	4	4	N	N	N	N	B15E0359	B15E0364	B15E0366	
A136(049)	S	3	N	3	4	4	N	N	N	N	B15E0342	B15E0333	B15E0335	
A136(050)	T	3	N	1	2	3	N	N	N	N	B15E0283	B15E0287	B15E0312	
A136(051)	S	3	N	1	4	4	N	N	N	N	B15E1905	B15E0268	B15E0272	
A136(052)	S	4	N	1	4	3	N	N	N	N	B15E0231	B15E0243		

Appendix B

Assessment Report - T6775

T6775-BCN			LEFT CIRCUIT					RIGHT CIRCUIT							
Tower Details			Insulators And Fittings					Insulators And Fittings					Defect Image Hyperlinks		
													1	2	3	4
Tower No	Tension or Suspension (T/S)	Paintwork	Insulators - Electrical (132kV)	Insulators - Electrical (66kV)	Insulators - Mechanical (Rust)	U bolts or Tower attachments		Insulators - Electrical (132kV)	Insulators - Electrical (66kV)	Insulators - Mechanical (Rust)	U bolts or Tower attachments					
BCN001	T	2	N	1	5	3		N	1	5	3					
BCN002	T	2	N	1	5	3		N	1	5	3					
BCN003	S	2	N	2	5	3		N	2	5	3					
BCN004	S	3	N	2	5	3		N	2	5	3					
BCN005	T	2	N	1	5	3		N	1	5	3					
BCN006	S	2	N	2	5	3		N	2	5	3					
BCN007	S	2	N	2	5	3		N	2	5	3					
BCN008	T	N	N	1	5	3		N	1	5	3					
BCN009	T	4	N	1	5	3		N	1	5	3					
BCN010	S	3	N	2	5	3		N	2	5	3					
BCN011	T	2	N	1	5	3		N	1	5	3					
BCN012	S	2	N	2	5	3		N	2	5	3		B10K0564			
BCN013	T	N	N	1	1	1		N	1	1	1					
BCN029	T	N	N	1	1	3		N	1	1	3					
BCN030	S	N	N	2	4	2		N	2	4	2					
BCN031	S	N	N	2	3	2		N	1	3	2					
BCN032	S	N	N	2	2	2		N	2	2	2					
BCN033	S	N	N	2	2	2		N	2	2	2					
BCN034	S	N	N	2	2	2		N	1	2	2					
BCN035	T	N	N	1	1	3		N	1	1	3					
BCN036	S	N	N	2	2	2		N	2	2	2					
BCN037	S	N	N	2	2	2		N	2	2	2		B10K1041			