AdeNet: Deep learning architecture that identifies damaged electrical insulators in power lines

Ademola Okerinde
Kansas State University
Manhattan, USA
okerinde@k-state.edu

Lior Shamir
Kansas State University
Manhattan, USA
lshamir@k-state.edu

William Hsu
Kansas State University
Manhattan, USA
bhsu@k-state.edu

Tom Theis
Kansas State University
Manhattan, USA
theis@k-state.edu

ABSTRACT

Ceramic insulators are important to electronic systems, designed and installed to protect humans from the danger of high voltage electric current. However, insulators are not immortal, and natural deterioration can gradually damage them. Therefore, the condition of insulators must be continually monitored, which is normally done using UAVs. UAVs collect many images of insulators, and these images are then analyzed to identify those that are damaged. Here we describe AdeNet as a deep neural network designed to identify damaged insulators, and test multiple approaches to automatic analysis of the condition of insulators. Several deep neural networks were tested, as were shallow learning methods. The best results (88.8%) were achieved using AdeNet without transfer learning. AdeNet also reduced the false negative rate to ~7%. While the method cannot fully replace human inspection, its high throughput can reduce the amount of labor required to monitor lines for damaged insulators and provide early warning to replace damaged insulators.

1 INTRODUCTION

Power-line insulators change over time because they are continuously exposed to the weather, including heat, sun, and moisture. Deterioration can eventually damage insulators so they no longer function. Therefore identifying damaged insulators is a critical safety task. However, changes in characteristics such as color do

not necessarily indicate that the insulator is nonfunctional. Therefore, identifying damaged insulators requires a diverse dataset of different types of damaged and undamaged insulators, at different ages and under different conditions. In this research, we explore various techniques to classify the condition of electrical insulators in power lines.

Automatically detecting anomalies in image data has a broad range of applications [2, 8]: automatically identifying product/equipment defects, medical diagnostics [31, 37], quality checks at factories, and routine maintenance procedures. Models have been developed that can provide accurate image data. While most deep learning frameworks classify images well, detecting outliers and anomalies is more challenging. Training data for such tasks is often unbalanced because the number of anomaly samples, in many cases, is lower than the number of regular samples. Additionally, in some cases, the distinguishing characteristics of regular samples and anomalies are subtle; cracks, flashovers, scratches, and the like are difficult to identify. An effective anomaly detection model must, therefore, learn to identify these details as distinguishing characteristics.

We developed a deep learning architecture, named AdeNet, to identify damaged power line insulators without pre-training. AdeNet also eliminates data augmentation overhead as discussed in [17]. As a result, our model requires less computation and storage, making it more suitable for the computing capabilities of mobile/embedded devices such as unmanned aerial vehicles (UAV). Analysis on the device can be used to decide to acquire additional data while the UAV is still on its mission. AdeNet compares favorably with state-of-the-art deep learning (DL) architectures and shallow learning (SL) techniques. Our experimental results also show that the ROC-AUC score is not the best performance metric for an imbalanced dataset. A harmonic mean of recall and precision as expressed by the f1-score is a better choice.

2 RELATED WORK

Detecting anomalies or outliers in image data is important research for computer vision. This task is often characterized by a number of challenges, so it differs from typical supervised machine learning. One of these challenges is that anomaly samples are far fewer than regular samples.

Many methods have been proposed to address this challenge at both the data and algorithm levels. In [3], a comprehensive real-world dataset is curated for unsupervised detection of anomalies.

A common approach is to train the classifier solely on the normal samples with the hope that the architecture will be robust enough to capture the intrinsic characteristic features of the normal class and thus can identify abnormal characteristics by inference [23]. Notable state-of-the-art methods include AnoGAN [25, 26], L2 and SSIM Autoencoder [3, 4], CNN Feature Dictionary [20], GMM-Based Texture Inspection Model [5], and Variation Autoencoder

Because untreated damaged insulators are inherently dangerous, automatic monitoring of power line insulators have been substantially researched [19, 21]. A basic approach for detecting insulators in aerial images used morphology analysis with Otsu threshold followed by a support vector machine (SVM) classifier [35]. Another approach detected edges and corners represented through MultiScale MultiFeature descriptors [14]. Lattice detection has also been tested with good results in detecting anomalies in insulators in power lines [38].

In the specific context of automatically detecting damaged insulators, Structural Similarity [36] can determine the degree of distortion in damaged insulators from undamaged ones. However, as shown in our experiment, this approach works well if the distortion is digital or artificial, not the result of natural degradation. Circular GLOH-Like descriptors were used for detecting and classifying insulators [18]. A simple method based on shape and the distribution of brightness was proposed as a way to automatically identify cables damaged by lightning [12]. Common deep neural network architectures such as SSD [15] and ResNet [10] have also been used to identify damaged power lines in aerial images [21].

[18] proposed an effective and reliable method of detecting anomalies in insulators using a deep learning technique for aerial images. In the proposed deep detection approach, the single shot multibox detector (SSD), a powerful deep meta-architecture, using two-stage fine-tuning could automatically extract multi-level features from aerial images, thus obviating manual extraction. Inspired by transfer learning, a two-stage fine-tuning strategy was implemented with separate training sets. In the first stage, the basic insulator model was obtained by fine-tuning the COCO model with aerial images, including different types of insulators and various backgrounds. In the second stage, the basic model was fine-tuned using training sets with specific insulator types and specific situations. After the two-stage fine-tuning, the well-trained SSD model can directly and accurately identify the insulator in aerial images.

According to [16], the insulator is important to a transmission line, and detecting defects in insulators relies heavily on insulator position. Traditional methods of insulator recognition depend on color features and geometric features. These methods are influenced illumination and background, among other things, resulting in poor generalization. [16] proposed a method where insulators were recognized using deep learning algorithms. First, [16] constructed a training dataset that had three categories: insulator, background, and tower. Second, [16] initialized the convolution neural networks as a six-level network and adjusted training parameters to train the model. Finally, the trained model was used to predict the candidate insulator position. With the help of a non-maximum suppression algorithm and the line fitting method, [16] identified the exact location of an insulator.

3 PROPOSED APPROACH

Paradigms of automatic image classification can be broadly divided into DL and SL. DL often requires more difficult training and a large training set, but it has the advantage of being non-parametric. It is often achieves more accurate than SL. SL requires fewer training samples and is normally less prone to overfitting. However, image features from each image can be computationally expensive, and the resulting machine learning model is not always as accurate as models based on deep learning. In this research, we attempted both DL and SL to compare the two approaches to automatically detecting damaged power line insulators.

3.1 SHALLOW LEARNING

For shallow learning, we tested the Udat shallow learning image analysis tool [28], which implements the Wndchrm method [22, 30]. Udat works by first extracting a large set of 2841 numerical image content descriptors from raw pixels and transforming the raw pixels [27]. The numerical image content descriptors are shape, edges, textures, polynomial approximation, statistical distribution of pixel values, and others for a comprehensive numerical representation of the image [29]. Detailed information about the Wndchrm algorithm can be found in [22, 27, 29, 30].

Once the numerical image content descriptors are computed for all images, the values are then used for mature classification algorithms like Random Forest, Support Vector Machine, Gradient Boosting, and Naive Bayes.

3.2 DEEP LEARNING

In deep learning, multiple layers of mini-algorithms, called neurons, work together to draw complex conclusions. We chose the AdeNet architecture and compared it with several deep neural network architectures, from the basic LeNet-5 to more complex ResNet-101, VGG19, and MobileNetV2.

3.2.1 MobileNetV2. MobileNetV2 [24] is a mobile architecture based on Convolutional Neural Networks (CNNs). The basic building block is a bottleneck depth-separable convolution with residuals. Our implementation of MobileNetV2, taken from tensorflow-keras applications, has 398, 690 trainable, and 14,000 non-trainable parameters. The total size of the model is 5.7MB.

3.2.2 AdeNet. AdeNet is a deep learning architecture implemented with three layers of CNNs, each with batch normalization, maxpooling, and ReLU with no dropout. One fully connected layer comes before the softmax layer. Figure 1 shows the architecture. It contains the initial fully convolutional layer with 32 filters. We always used kernel size 3×3 as is standard for modern networks. The trained model size was 1.3MB with 102,082 trainable parameters and 448 non-trainable parameters.

4 METHODOLOGY AND EXPERIMENTAL DESIGN

4.1 Dataset

The dataset was provided by Black and Veatch, consisting of 1696 JPEG images of power-lines with the dimensions of 5280×3956 and a resolution of 72x72. We randomly divided the entire dataset into
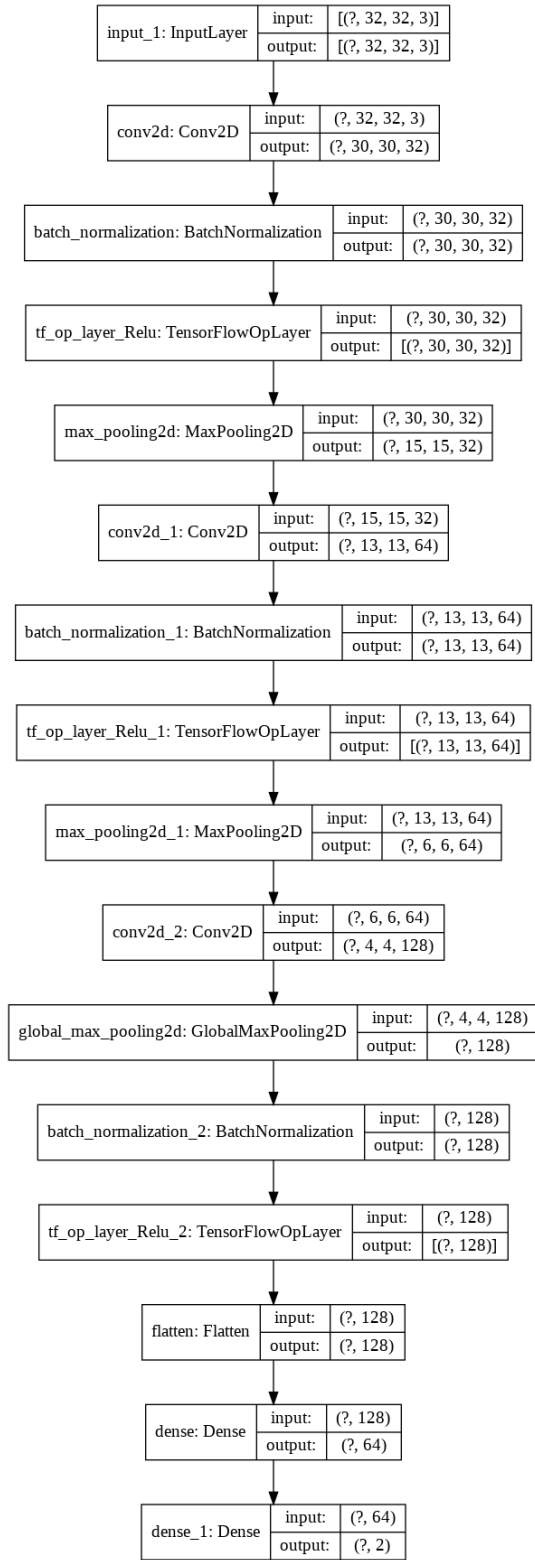
input_1: InputLayer	input:	[(?, 32, 32, 3)]
	output:	[(?, 32, 32, 3)]

conv2d: Conv2D	input:	(?, 32, 32, 3)
	output:	(?, 30, 30, 32)

batch_normalization: BatchNormalization	input:	(?, 30, 30, 32)
	output:	(?, 30, 30, 32)

tf_op_layer_Relu: TensorFlowOpLayer	input:	(?, 30, 30, 32)
	output:	[(?, 30, 30, 32)]

max_pooling2d: MaxPooling2D	input:	(?, 30, 30, 32)
	output:	(?, 15, 15, 32)

conv2d_1: Conv2D	input:	(?, 15, 15, 32)
	output:	(?, 13, 13, 64)

batch_normalization_1: BatchNormalization	input:	(?, 13, 13, 64)
	output:	(?, 13, 13, 64)

tf_op_layer_Relu_1: TensorFlowOpLayer	input:	(?, 13, 13, 64)
	output:	[(?, 13, 13, 64)]

max_pooling2d_1: MaxPooling2D	input:	(?, 13, 13, 64)
	output:	(?, 6, 6, 64)

conv2d_2: Conv2D	input:	(?, 6, 6, 64)
	output:	(?, 4, 4, 128)

global_max_pooling2d: GlobalMaxPooling2D	input:	(?, 4, 4, 128)
	output:	(?, 128)

batch_normalization_2: BatchNormalization	input:	(?, 128)
	output:	(?, 128)

tf_op_layer_Relu_2: TensorFlowOpLayer	input:	(?, 128)
	output:	[(?, 128)]

flatten: Flatten	input:	(?, 128)
	output:	(?, 128)

dense: Dense	input:	(?, 128)
	output:	(?, 64)

dense_1: Dense	input:	(?, 64)
	output:	(?, 2)

Figure 1: AdeNet Deep Learning Architecture

Figure 2: An example of a power-line image.

80% training and 20% test sets. Each image is annotated with a bounding-box around the objects of interest: the insulators. The insulators are broadly categorized as damaged (positive class) and undamaged (negative class). Table 1 shows the number of samples in each class after cropping out the insulators.

Table 1: Number of samples of damaged and undamaged insulators in the dataset.

Dataset	Damaged	Undamaged	#images
Train + Validation	1417	2836	1484
Test	290	835	212

The dataset presents a number of interesting challenges: within class variations, class imbalance, noisy background, and varying orientations and scales. Using within class variations as an example, in our experiments, we classified as damaged any insulator with flash-over or that was broken, missing, or fried. This leads to skewness of samples within each class as well.

4.2 Data Pre-processing

Figure 2 shows an example of a power-line image. The electrical insulators were then cropped out of images for feature extraction, giving us a total of 4253 images for training and validation as well as 1125 images for testing. Figure 3 shows samples of damaged insulators, and Figure 4 shows samples of undamaged insulators. For SL, we used Udat to extract the 2900 most informative image features from the insulators. Udat uses statistical techniques like Radon transform features, Chebyshev Statistics, and Multi-scale Histograms for image feature extraction. For DL, because insulators varied in scale, we padded them to the max dimension in each batch during training and testing.

5 RESULTS AND DISCUSSION

The different methods were applied to the dataset described in Section 4.1. For training and testing, we used mostly 5-fold as the
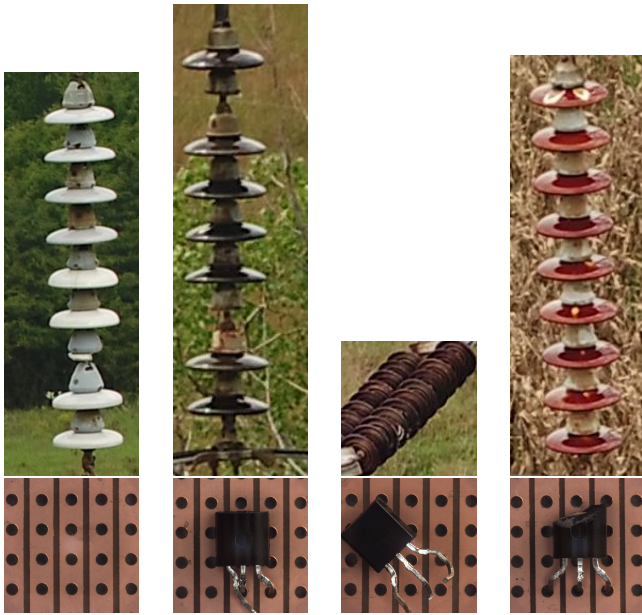
Figure 3: Damaged insulators (top) and transistors (bottom)

different classifiers for cross-validation. Each split was trained with 10 or 20 epochs. For transfer learning, we used models pre-trained on the ImageNet dataset. We used a batch size of 16 for training and testing. Most of our DL experiments were run on Tesla P100-PCIE on Google Colab. The DL models were built using TensorFlow [1].

The results are shown in tables 7 and 8. For the analysis of the results, the table gives accuracy, F1 score, precision, recall, and ROC area under the curve for each classifier. The measurements are the average of all folds.

We found that MobileNetV2 sometimes performed better if trained using more epochs, starting with saved model weights. The performance often dropped if the F1-score for the initial 10 epochs reached 91% in the damaged class. The model often inverted performance from one class to favor another if the results for the first 10 epochs were skewed. As tables 3 and 4 show, the model had fewer false negatives when trained for fewer epochs. This is crucial for this situation.

Without transfer learning, MobileNetV2 was highly influenced by the class imbalance because most data were classified as negatives. Likewise, LeNet-5 [13], EfficientNetB7 [33], VGG19 [32], and ResNet-101 [10] models all classified most data as negatives. All models other than LeNet-5 were pre-trained on the ImageNet dataset [6]. Table 7 shows their performance on different validation percentage splits.

Our experimental findings showed that all MobileNetV2 model performance on the test set correlated with results from the validation set. For example, if a model did well on the positive class in validation, it also performed well on the positive class in the test and vice versa.

I'm not sure what this means. Technically, this says that if a model performed well in the test, it also performed well on the positive class. Is that correct? Or do you mean that if a model did well on the negative

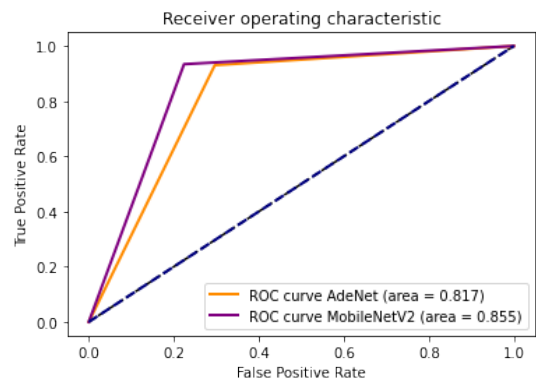**Figure 4: Undamaged insulators (top) and transistors (bottom)**

Figure 5: Model comparison of the ROC curve on insulator dataset

Figure 6: Grad-CAM heatmap visualizations

class, it did well in the test? Or does it mean that if the model did poorly in the positive class, it did poorly in the test. Basically, vice versa does not work here. State it clearly.

We ran all the five models on the test set, and the results showed that MobileNetV2 performed better on 20-epochs-models on the test set. We reported the average of the five models in Table 8.

Comparing AdeNet with similar simple DL architectures like LeNet5 revealed that an overly simple network was not optimal, but a carefully engineered simple architecture that balanced between simple and complex designs, like ResNet-101 and VGG19, worked. The confusion matrices in tables 5, 6, 4, 3 and 2

Put the tables in numerical order.

reveal that AdeNet was significantly better at classifying both classes of the dataset. This was confirmed by high f1 scores in table 8. F1 score is a better performance metric for evaluating models developed for imbalanced datasets. Consider, for instance, random forest with high roc auc

This abbreviation (roc auc) was different earlier. Keep everything consistent throughout.

score, but poor f1 score (see table 2).

The small number of samples in [3]'s transistor category means AdeNet performed poorly with and without fine-tuning because the entire dataset was classified as negatives, i.e., undamaged transistors. Borrowing intuition from [34], we experimented with freezing two, ten, and thirteen layers of pre-trained AdeNet on the insulator dataset to see how well it transferred to the transistor detection task. This reinforced the conclusion that a DL model performs better with more data. However, SL with Udat achieved an accuracy of 92.8% (10-fold-cross-validation split) on validation data with roughly half of the damaged class miss-classified.

With Grad-CAM, we used heatmap to visualize the model's attention on the insulator while making predictions as shown in Figure 6. Grad-CAM uses the penultimate (pre-dense) convolutional layer containing spatial information that is completely lost in dense layers. The model's reasoning did make some sense from the human visual system perspective. The area that constituted damage in the insulators were located where the heatmap was most intense, i.e., heavily red.

Figure 5 compares the ROC curve for the models using the Black and Veatch insulator dataset. The curve shows that MobileNetV2 was slightly better than AdeNet. This plot was based on one model. Table 8 shows that, on average, AdeNet outperformed MobileNetV2.

A major hyper-parameter tuning task was to determine how long to train a deep neural network. We experimented with using early stopping and fixed epochs, finding that performance was much the same. However, with callbacks, training time was significantly reduced.

Of the tested classifiers, the best shallow learning classifier was Random Forest. However, its confusion matrix (see Table 2) revealed that the classifier did not do well in classifying the damaged class, making it the worst classifier for this task.

5.1 Ablation Studies

We inspected parts of the architecture and parameters for analysis.

Batch Normalization We observed no significant differences in model performance with or without batch normalization [11]

Table 2: Confusion matrix for Random Forest on test set

	Predicted Damaged	Predicted Undamaged
Actually Damaged	11	279
Actually Undamaged	2	833

Table 3: Confusion matrix for MobileNetV2 after 20 epochs for test set

	Predicted Damaged	Predicted Undamaged
Actually Damaged	991	459
Actually Undamaged	1063	3112

Table 4: Confusion matrix for MobileNetV2 after 10 epochs for test set

	Predicted Damaged	Predicted Undamaged
Actually Damaged	724	726
Actually Undamaged	312	3863

Table 5: Confusion matrix for AdeNet after 10 epochs for test set

	Predicted Damaged	Predicted Undamaged
Actually Damaged	947	503
Actually Undamaged	257	3918

possibly because AdeNet is not exceptionally deep. We could thus further reduce the parameter by ~1%.

6 CONCLUSION

In this research, we addressed the task of detecting damaged electrical insulators in power-line images, an important safety task. Detecting damaged electrical insulators has been a labor-intensive task requiring substantial experience and careful examination of the images.

The new architecture, AdeNet, based on a deep convolutional neural network has the advantage of requiring little energy, allowing it to be used on low-energy devices like UAVs, which are often used to acquire the data for identifying damaged insulators. The low energy allows AdeNet to make decision on the UAV, which can then be used to acquire more data of suspect insulators during a mission, without going back to a site and re-using the UAV. Like intuition in [7, 9], we found an architecture that is not as simple as LeNet5 and not as complex as ResNet (just the right mix of simplicity and complexity) that also performs well in detecting anomalies in damaged insulators in power lines.

Check for consistency in the phrase power lines. I've seen it punctuated different ways. It should be the same throughout.

In comparing the proposed method and other solutions, including shallow learning, we found that deep learning outperforms shallow learning architectures. Within DL architecture, fine-tuning the pre-trained CNN models did not improve the performance in all cases. We demonstrated this through experiments with our AdeNet

Table 6: Confusion matrix for AdeNet after 20 epochs for test set

	Predicted Damaged	Predicted Undamaged
Actually Damaged	1026	424
Actually Undamaged	213	3962

architecture. Experimental results also showed that our architecture, without pre-training, outperformed all other deep neural networks. This makes AdeNet suitable for mobile/embedded devices that have storage and computation constraints.

ACKNOWLEDGMENTS

REFERENCES

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467* (2016).

[2] Laura Beggel, Michael Pfeiffer, and Bernd Bischl. 2019. Robust anomaly detection in images using adversarial autoencoders. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 206–222.

[3] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. 2019. MVTec AD–A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 9592–9600.

[4] Paul Bergmann, Sindy Löwe, Michael Fauser, David Sattlegger, and Carsten Steger. 2018. Improving unsupervised defect segmentation by applying structural similarity to autoencoders. *arXiv preprint arXiv:1807.02011* (2018).

[5] Tobias Böttger and Markus Ulrich. 2016. Real-time texture error detection on textured surfaces with compressed sensing. *Pattern Recognition and Image Analysis* 26, 1 (2016), 88–94.

[6] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 248–255.

[7] Jonathan Frankle and Michael Carbin. 2018. The lottery ticket hypothesis: Finding sparse, trainable neural networks. *arXiv preprint arXiv:1803.03635* (2018).

[8] Izhak Golan and Ran El-Yaniv. 2018. Deep anomaly detection using geometric transformations. In *Advances in Neural Information Processing Systems*. 9758–9769.

[9] Guy Hacohen and Daphna Weinshall. 2019. On the power of curriculum learning in training deep networks. *arXiv preprint arXiv:1904.03626* (2019).

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.

[11] Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167* (2015).

[12] R Ishino and F Tsutsumi. 2004. Detection system of damaged cables using video obtained from an aerial inspection of transmission lines. In *IEEE Power Engineering Society General Meeting, 2004*. IEEE, 1857–1862.

[13] Yann LeCun et al. 2015. LeNet-5, convolutional neural networks. *URL: http://yann. lecun. com/exdb/lenet* 20, 5 (2015), 14.

[14] Shenglong Liao and Jubai An. 2014. A robust insulator detection algorithm based on local features and spatial orders for aerial images. *IEEE Geoscience and Remote Sensing Letters* 12, 5 (2014), 963–967.

[15] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. 2016. Ssd: Single shot multibox detector. In *European Conference on Computer Vision*. Springer, 21–37.

[16] Yue Liu, Jun Yong, Liang Liu, Jinlong Zhao, and Zongyu Li. 2016. The method of insulator recognition based on deep learning. In *2016 4th International Conference on Applied Robotics for the Power Industry (CARPI)*. IEEE, 1–5.

[17] Kamal Malik, Harsh Sadawarti, and Kalra G S. 2014. Comparative analysis of outlier detection techniques. In *IJCA*, Vol. 97. 12–21.

[18] Xiren Miao, Xinyu Liu, Jing Chen, Shengbin Zhuang, Jianwei Fan, and Hao Jiang. 2019. Insulator detection in aerial images for transmission line inspection using single shot multibox detector. *IEEE Access* 7 (2019), 9945–9956.

Table 7: Classification accuracy, F1, precision, recall, and ROC area under the curve when using different methods on varying validation set.

Dataset	Classifiers	Acc	F1	Precision	Recall	ROC Area	#folds
val	Random Forest	0.86	0.85	0.86	0.86	0.92	10
val	MultiLayer Perceptron	0.82	0.82	0.82	0.82	0.86	10
val	Gradient Boosting	0.87	-	-	-	-	
val	LeNet-5	0.68	0.46	0.70	0.53	0.53	5
val	AdeNet + 10 epochs	0.86	0.83	0.87	0.81	0.81	5
val	EfficientNetB7	0.65	0.55	0.57	0.55	0.55	1
val	VGG19	0.67	0.40	0.44	0.50	0.50	5
val	ResNet-101	0.59	0.40	0.35	0.51	0.51	5

Table 8: Classification accuracy, F1, precision, recall, and ROC area under the curve when using different methods on test data (average of 5 folds cross validation).

Learning	Classifiers	Acc	F1	Precision	Recall	ROC Area
Shallow	Random Forest	0.79	0.64	0.79	0.62	0.82
Shallow	Random Tree	0.69	0.60	0.60	0.60	0.60
Shallow	Naive Bayes	0.56	0.56	0.64	0.67	0.76
Shallow	MultiLayer Perceptron	0.77	0.69	0.69	0.68	0.79
Shallow	Support Vector Machine	0.72	0.65	0.65	0.66	0.66
Deep	MobileNetV2 + 10 epochs	0.82	0.72	0.77	0.71	0.73
Deep	MobileNetv2 + 20 epochs	0.70	0.63	0.76	0.70	0.71
Deep	AdeNet + 10 epochs	0.86	0.81	0.85	0.80	0.80
Deep	**AdeNet + 20 epochs**	**0.89**	**0.84**	**0.87**	**0.83**	**0.83**

[19] François Mirallès, Nicolas Pouliot, and Serge Montambault. 2014. State-of-the-art review of computer vision for the management of power transmission lines. In *Proceedings of the 2014 3rd International Conference on Applied Robotics for the Power Industry*. IEEE, 1–6.

[20] Paolo Napoletano, Flavio Piccoli, and Raimondo Schettini. 2018. Anomaly detection in nanofibrous materials by cnn-based self-similarity. *Sensors* 18, 1 (2018), 209.

[21] Van Nhan Nguyen, Robert Jenssen, and Davide Roverso. 2018. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *International Journal of Electrical Power & Energy Systems* 99 (2018), 107–120.

[22] Nikita Orlov, Lior Shamir, Tomasz Macura, Josiah Johnston, D Mark Eckley, and Ilya G Goldberg. 2008. WND-CHARM: Multi-purpose image classification using compound image transforms. *Pattern Recognition Letters* 29, 11 (2008), 1684–1693.

[23] Marco Rudolph, Bastian Wandt, and Bodo Rosenhahn. 2020. Same Same But DifferNet: Semi-Supervised Defect Detection with Normalizing Flows. *arXiv preprint arXiv:2008.12577* (2020).

[24] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4510–4520.

[25] Thomas Schlegl, Philipp Seebock, Sebastian M Waldstein, Georg Langs, and Ursula Schmidt-Erfurth. 2019. f-anogan: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Analysis* 54 (2019), 30–44.

[26] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. 2017. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International Conference on Information Processing in Medical Imaging*. Springer, 146–157.

[27] Lior Shamir. 2008. Evaluation of face datasets as tools for assessing the performance of face recognition methods. *International Journal of Computer Vision* 79, 3 (2008), 225.

[28] Lior Shamir. 2017. UDAT: A multi-purpose data analysis tool. *ASCL* (2017), ASCL–1704.

[29] Lior Shamir, Tomasz Macura, Nikita Orlov, D Mark Eckley, and Ilya G Goldberg. 2010. Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. *ACM Transactions on Applied Perception* 7, 2 (2010), 1–17.

[30] Lior Shamir, Nikita Orlov, D Mark Eckley, Tomasz Macura, Josiah Johnston, and Ilya G Goldberg. 2008. Wndchrm–an open source utility for biological image analysis. *Source code for biology and medicine* 3, 1 (2008), 13.

[31] Desire Sidibe, Shrinivasan Sankar, Guillaume Lemaitre, Mojdeh Rastgoo, Joan Massich, Carol Y Cheung, Gavin SW Tan, Dan Milea, Ecosse Lamoureux, Tien Y Wong, et al. 2017. An anomaly detection approach for the identification of DME patients using spectral domain optical coherence tomography images. *Computer Methods and Programs in Biomedicine* 139 (2017), 109–117.

[32] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[33] Mingxing Tan and Quoc V Le. 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946* (2019).

[34] Xin Wang, Thomas E Huang, Trevor Darrell, Joseph E Gonzalez, and Fisher Yu. 2020. Frustratingly Simple Few-Shot Object Detection. *arXiv preprint arXiv:2003.06957* (2020).

[35] Xian Wang and Youmin Zhang. 2016. Insulator identification from aerial images using support vector machine with background suppression. In *International Conference on Unmanned Aircraft Systems*. IEEE, 892–897.

[36] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* 13, 4 (2004), 600–612.

[37] Qi Wei, Yinhao Ren, Rui Hou, Bibo Shi, Joseph Y Lo, and Lawrence Carin. 2018. Anomaly detection for medical images based on a one-class classification. In *Medical Imaging 2018: Computer-Aided Diagnosis*, Vol. 10575. International Society for Optics and Photonics, 105751M.

[38] Jingjing Zhao, Xingtong Liu, Jixiang Sun, and Lin Lei. 2012. Detecting insulators in the image of overhead transmission lines. In *International conference on intelligent computing*. Springer, 442–450.