Embedded Self Organizing Systems (Vol 10. No 5. 2023) (pp.13-17)

Embedded Selforganizing Systems

Exploring SSD Detector for Power Line Insulator Detection on Edge Platform

Mohamed Salim Harras
Department of Computer Engineering
Chemnitz University of Technology
Chemnitz, Germany
mohamed-salim.harras@informatik.tu-chemnitz.de

Soaibuzzaman
Department of Software Engineering
Bauhaus-Universität Weimar
Weimar, Germany
soaibuzzaman@uni-weimar.de

Shadi Saleh
Department of Computer Engineering
Chemnitz University of Technology
Chemnitz, Germany
shadi.saleh@informatik.tu-chemnitz.de

Wolfram Hardt
Department of Computer Engineering
Chemnitz University of Technology
Chemnitz, Germany
wolfram.hardt@informatik.tu-chemnitz.de

***Abstract*—Power line insulator detection is pivotal for the consistent performance of the entire power system. It forms the basis of Unmanned Aerial Vehicle (UAV) inspection, an emerging trend in power line surveillance. This paper addresses the challenge of insulator detection in cluttered aerial images, given the constraints of a limited dataset and lower computational resources, specifically on the NVIDIA Jetson Nano platform. We have developed two approaches based on active and passive deep learning algorithms, underpinned by the Single Shot Multibox Detector (SSD) meta-architecture with MobileNetV2 as its backbone - SSD300 and SSD640. The proposal models managed a frame rate of 9 fps in 10W power mode and 5.6 fps in 5W power mode. Our experiments demonstrated that the proposed active learning model could conduct robust insulator detection, achieving a mAP of 94.5% while using only 43% of the total dataset, comparable to the traditional deep learning approach's 94.6% mAP using the entire dataset. Significantly, the active learning model seeks feedback during the training process, allowing the model to learn from its mistakes while enhancing its accuracy over time. This also contributes to improved generalizability and interpretability of the model by seeking diverse and representative samples during training, all while reducing the computational and annotation overhead.**

Keywords— Insulator detection, single shot detector, deep learning, active learning, embedded platforms

I. Introduction

Electricity, a cornerstone of modern civilization, requires an efficient and reliable delivery system. At the heart of this system, the transmission network performs a crucial function by ferrying power from generation stations to local distribution networks. Such a network's robustness directly impacts the overall reliability and performance of the entire power system. Within the components of high-voltage transmission lines, insulators play a vital role. They not only maintain electrical insulation but also provide mechanical support to conductors. Thus, the detection and inspection of power line insulators become a critical task in preserving the consistent operation of power systems.

In the past, the methods for inspecting insulators primarily involved the use of binoculars or helicopters. However, recent years have witnessed a growing trend toward using Unmanned Aerial Vehicles (UAVs) for these inspection tasks. This shift is attributable not only to the unique airborne perspective these vehicles provide but also to their ability to closely approach a power line. At the same time, electricity continues to flow, thus increasing efficiency and safety.

The meticulous examination of power line insulators for defects is essential to ensure a stable power grid. Nonetheless, before any inspection can occur, there is a prerequisite need for detecting the insulator. This task has been the subject of many studies, leading to the development of various methods aimed at effective insulator detection. This task can be addressed with camera sensors through a computer vision application. Traditional computer vision tasks, however, are known for their sensitivity to multiple factors, underscoring the importance of deep learning models. With their robust capabilities, these models present a more reliable solution for the detection of objects, providing a promising avenue for the future of power line insulator detection.

II. Literature Review

In recent years, special attention has been given toward the automation of power line insulator inspections. This was guided by the desire to reduce the risks and costs related to this process. The inspection process requires first the correct and stable detection of the insulators.

A. High-Voltage Insulator Detection

Earlier insulator detection strategies primarily employed traditional image processing and computer vision techniques. These methods offer less computational complexity, power, and resource consumption. For example, Fei et al. [1]utilized Bayesian segmentation for insulator detection, while Zhai et al.[2] employed adaptive morphology and saliency to identify faulty insulators. Another effective approach is Li et al.[3]'s usage of template matching to detect insulators from UAV images. Oberweger et al. [4]proposed a RANSAC-based model for insulator detection that used a voting approach. Other noteworthy methodologies include feature-based algorithms like those proposed by Liao and An[5], as well as Tudevdagva et al.[6], who focused on detecting faults in insulators based on symmetry detection.

Recently, more focus has been invested in AI-based methods to solve the detection task. With the progress of deep learning and CNNs, precise methods for insulator detection have been proposed, improving upon traditional methods in terms of accuracy and speed. These models have been commonly classified into two categories: single-stage detectors and two-stages detectors[7].

For instance, Zhao et al.[8] proposes a model for insulator recognition and fault detection in transmission lines, using an enhanced Faster RCNN model with Feature Pyramid Networks and various image processing techniques. Their model demonstrated high accuracies exceeding 90% in mAP. Another research [9] outlines a novel, robust approach using deep learning to detect cracked insulators in high voltage transmission lines. A Region-based Fully Convolutional Networks (R-FCN) algorithm was used to localize the insulator followed by the inspection step. The model demonstrates an average accuracy rate of 90.5% and strong environmental adaptability.

Han et al. [10]presents an improved insulator defect detection algorithm using YOLOX. An enhanced SIoU loss function is used to address issues related to minor target sensitivity and complex backgrounds in aerial insulator images. By incorporating an Efficient Channel Attention Module (ECA), the model reduces the impact of redundant features on detection accuracy, achieving an impressive mAP of 97.18% and a detection speed of 71 fps. Another study [11] used an improved SSD algorithm for real-time UAV-based power inspection, utilizing a lightweight MnasNet network and multiscale feature fusion technique, achieving a detection accuracy of 93.8% and a speed of 154 ms per frame.

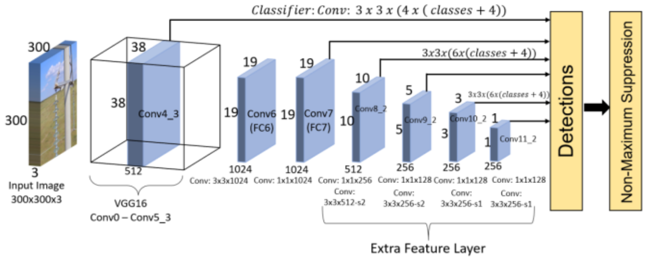
In a previous study[12], we trained a YOLOv4-Tiny model and analysed the performance in the scope of an embedded platform. The results showcased that real-time detection was possible thanks to the lightweight model and a specific power configuration of the embedded board.

B. Passive vs. Active Learning

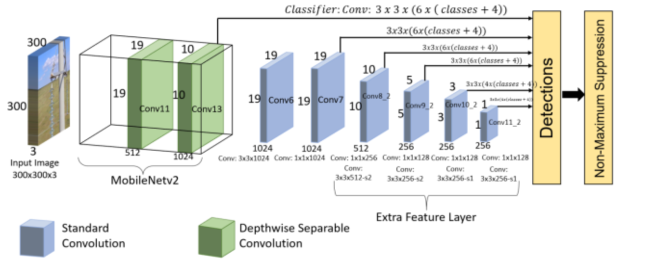
The application of deep learning in visual inspection systems has surged in the past decade due to advances in neural networks, greatly enhancing their robustness. The quality of these models relies heavily on the architecture of the deep learning and the strategies used in training.

Traditional supervised learning, in other words passive learning, refers to the process where a model is trained on a labelled dataset without any interaction during the training phase. This learning process is referred to as passive since it learns independently from the data without any guidance. This means that once the model is fed with data, it would train it on one iteration. Therefore, it highly relies on the quality of the dataset to produce accurate models.

However, the approach of passive learning presents several limitations, leading to the rise of active learning. Contrasting passive learning, which relies on a fixed dataset with pre-labelled examples, active learning actively selects samples for labelling to optimize its learning process. Active learning offers the advantage of reducing the annotation cost by prioritizing the most informative instances, while passive learning follows a more traditional approach of training on a fixed labelled dataset. The choice between passive and active learning for inspection problems depends on the availability of labelled data, the cost of annotation, and the desired efficiency and accuracy trade-offs.



(a)



(b)

Fig. 1 Network architecture of SSD (a) with standard VGG16 backbone (b) with proposed MobileNetv2

Active learning is a machine learning methodology designed to augment model accuracy while reducing the need for labelled data. A semi-supervised learning method selects the most informative examples from a large unlabelled data set and presents them to a human expert for labelling. After these examples are labelled, they are used to train the model in an iterative process. According to Shadi et al.[13], [14], active learning can greatly decrease the need for labelled data and match, or even exceed, the accuracy of passive deep learning methods.

In visual inspection tasks, where labelled data is scarce or expensive to obtain, active learning can be a useful approach. It offers the benefit of minimizing the need for data annotation and can achieve performance on par with deep passive learning methods but with fewer labelled datasets However, passive deep learning could be a viable choice if there is easy access to a large and labelled dataset and the cost of annotation is not a significant issue. The unique constraints, available resources, and preferred compromises regarding data efficiency and annotation costs dictate the decision between these two methods [14].

III. METHODOLOGY

This study introduces a methodology to automate the insulator detection process using a SSD architecture. This approach involves a process that includes data collection, data augmentation, and training based on both traditional and active learning strategies.

The SSD consists of an object detector that performs the detection in one forward pass through the network, contrasting other models such as RCNN and its variants[15]. The SSD model categorizes the output area of bounding boxes into a collection of default boxes. This is achieved by altering their scales and aspect ratios for each location on the feature map. During the prediction phase, the network generates probability scores for each category in each default box and then adapts the shape of the box to fir the object. In order to handle the detection of object with different sizes, the model combines the predictions of feature maps at different resolutions[15].

In the context of this research, we modified the backbone of the SSD model by integrating the MobileNetv2 network instead of the VGG16. This modification reduces the computational complexity, facilitating feature extraction and enabling faster training times, which is beneficial for an iterative training process like active learning. In addition, it reduces the number of network parameters, making the model compact and more suitable for resource-limited edge devices. The model architecture is shown in Fig 1.

A. Dataset Collection

In this work's context, the dataset used in [12] was reutilized to save the labelling cost for this research work [12]. However, it was necessary to revise the dataset quality. The dataset's quality is crucial in ensuring a high-calibre model's production. Empirical evidence demonstrates a direct correlation between the dataset's quality and the resultant model's performance and accuracy [16]

B. Data Augmentation

Data augmentation is a technique that allows for the expansion of feature variety in the data set used for model training[17]. This technique amplifies a model's ability to fortify its resistance to changes in object positioning and mitigates overfitting. To streamline the training duration and reduce resource use, we favoured offline augmentation over the online method. We implemented several augmentation techniques such as:

- Brightness (±20%)
- Exposure (±20%)
- Rotation
- Flipping
- Noise (salt and pepper)

In the context of this work and given that we planned to train our model using two different resolutions, namely 640x640 (SSD640) and 300x300 (SSD300), we initially extracted frames at a 640x640 resolution. Subsequently, these images were transformed into a resolution of 300x300 to form a separate dataset. This procedure led to the assembly of a dataset comprising 2,337 images, which we divided as explained in the methodology.

C. Deep Learning Frameworks

Our research incorporates two distinguished deep learning frameworks: conventional learning and active learning. Each learning framework brings specific benefits and challenges to the table. These learning approaches determine how our model engages with the training data, ultimately influencing the model's performance and reducing the effort required in the training phase.

1) Passive Learning

In the conventional supervised learning approach, also known as passive deep learning, the model is trained once on a fully labelled dataset without further interaction or adjustment during the learning process. This is considered an unsupervised learning process. For this study, the dataset was divided into an 80:20 ratio, with 80% allocated for training and the remaining 20% for testing purposes. The SSD model was trained and subsequently fine-tuned using a validation set comprising around 20% of the training set.

2) Active Learning

Active learning consists of a semi-supervised technique that focuses on iteratively selecting the most informative instances from a pool of unlabelled data for manual labelling, decreasing the volume of labelled data necessary for effective training. The active deep learning methodology utilized for the detection of insulators encompasses several stages as depicted in Fig 2.

The procedural steps within the proposed framework are as follows:

1. Generate a data pool containing all chosen images derived from the aerial video footage.
2. To initiate, select 15% of the total images for the first iteration. The data pool comprises 7008 images, and 1051 (15%) are selected for the first iteration.
3. Use an oracle (such as a human expert) to annotate the chosen dataset.
4. Train the proposed modified SSD model using the annotated dataset.
5. Post-training, verify if the model achieves the required mAP accuracy or if the loss function has stabilized. If these conditions are not met, draw an additional 4% of images from the remaining data pool.
6. Until the model meets the required mAP accuracy or the loss function stabilizes, repeat steps 3 and 4 as needed.
7. The final model exhibiting the highest mAP is chosen for inference.

The strategy for selecting queries is a crucial element of the active learning paradigm. Selecting an apt query strategy is crucial for achieving optimal results in minimal time. As mentioned in the active learning process, the data is collected from data pools containing data organized based on a specific logic. In the context of this work, the dataset was split into four pools based on several features such as light, reflection, blur, noise, shadow, angle, and complexity of background. The four pools are as follows:

- High-quality Pool: This category comprises images exhibiting the maximum confidence score. These images are free from any distortions, such as reflection, blur, noise, or shadows, and the background is free from clouds or artifacts. Moreover, these images were captured in proximity to the insulator.

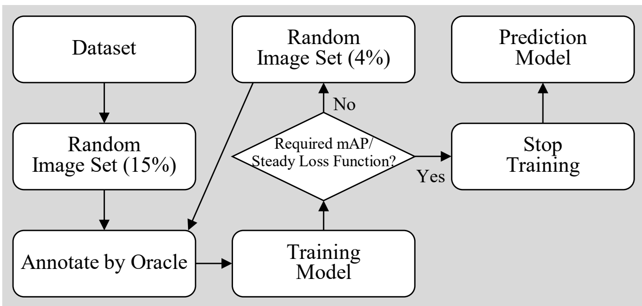


Fig. 2 Proposed framework for active deep learning

- Intermediate-quality Pool: This category excludes images with reflections, blur, or noise. However, it includes images that are tilted or partially obstructed. The background may occasionally display clouds but lacks any additional artifacts.
- Low-quality Pool: This category includes images that might be periodically out of focus, resulting in blurry depictions of the insulators. Some e images exhibit lens flares, and the backgrounds can occasionally be chaotic, with visible houses or traffic.
- Indoor Pool: This category consists of images captured within the Indoor Flight Centre (IFC) lab at TU Chemnitz. These images are shot under various indoor lighting conditions instead of natural daylight. While the background remains relatively static, the images are captured from various angles.

D. Model Training

The model was trained both on passive learning and active learning. For passive learning, a batch size of 32 was used to train a model on 10000 epochs. The same hyperparameters were chosen for the two types of learning paradigms.

IV. Results and Discussion

This section discusses the results and evaluation of our trained models. Our training process encompassed two models: SSD300 and SSD640, employing both passive and active learning techniques. We proceeded to assess the performance of these models using evaluation metrics based on the confusion matrix. Finally, we deployed the models on the NVIDIA Jetson Nano platform and measured their detection speed and inference rate, providing valuable insights into their real-time performance.

Table I presents the evaluation results of SSD300. For mAP@0.50, both the passive and active learning (8th iteration) approaches achieved comparable precision, with 86.7% and 83.7%, respectively. However, when considering the accuracy of the test set, the active approach outperformed the passive approach significantly, achieving 71.2% (on the 8th iteration) accuracy compared to only 52.3% in the passive approach.

The evaluation results of SSD640 are presented in Table II. After eight iterations, both approaches demonstrate nearly identical mean average precision values for both mAP@0.50:0.95 and mAP@0.50. In contrast to SSD300, the passive approach performs slightly better on the test set, achieving 80.7% compared to the 77.9% accuracy on the 8th iteration, during the 8th iteration, the active deep learning (DL) approach only utilized 43% of the total dataset, highlighting its efficiency in leveraging a smaller portion of the data for achieving promising results.

TABLE I. Evaluation results of SSD300

Strategy	mAP@0.5 (%)	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
Passive Learning	86.7	52.3	99.7	68.6	52.3
Active Learning Iterations	76.5	53.8	98.9	69.7	53.4
	78.4	59.6	98.8	74.4	59.2
	81.4	59.62	98.8	74.4	59.2
	81.6	58.66	1.0	73.9	58.7
	83.0	64.0	99.5	77.9	63.8
	83.3	65.8	98.2	78.8	64.9
	82.3	71.5	98.8	83.0	70.9
	83.7	71.6	99.3	83.2	71.2

TABLE II - Evaluation results of SSD640

Strategy	mAP@0.5 (%)	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
Passive Learning	94.6	89.8	88.8	89.3	80.7
Active Learning Iterations	82.2	86.6	77.3	81.6	69.0
	85.1	87.5	81.8	84.5	73.2
	86.1	86.3	80.9	83.5	71.7
	87.7	89.4	79.7	84.3	72.8
	89.1	86.1	84.1	85.1	74.0
	91.4	93.3	81.4	86.9	76.9
	91.6	94.9	81.4	87.6	77.9
	94.5	95.2	81.1	87.6	77.9

In addition to evaluating the model's performance, we also measure the inference time to assess its efficiency. Subsequently, the frame rate is derived from the recorded time, indicating the number of frames processed by the model per second in a video streaming scenario. By calculating the Frames Per Second (FPS) as the inverse of the inference time (FPS = 1/inference time), we gain insight into the model's ability to meet real-time constraints. Results of the model's detection can be seen in Figure 3.

Table III comprehensively compares the frame rate and average power consumption of the NVIDIA Jetson Nano. The inference speed is evaluated under two different power modes: 5W and 10W, available on the Jetson Nano.

In the case of SSD300, when operating at 5W power mode, the model trained using the passive deep learning strategy achieves an average frame rate of 4.2 fps, whereas, at 10W power mode, it reaches 9.0 fps. Comparatively, the active DL approach yields slightly higher inference rates and consumes slightly less power than the passive strategy.

With the increase in image size for SSD640, the frame rate experiences a significant decrease, nearly halving its value. When employing the passive DL approach, the frame rates observed are 2.2 fps in the 5W power mode and 3.4 fps in the

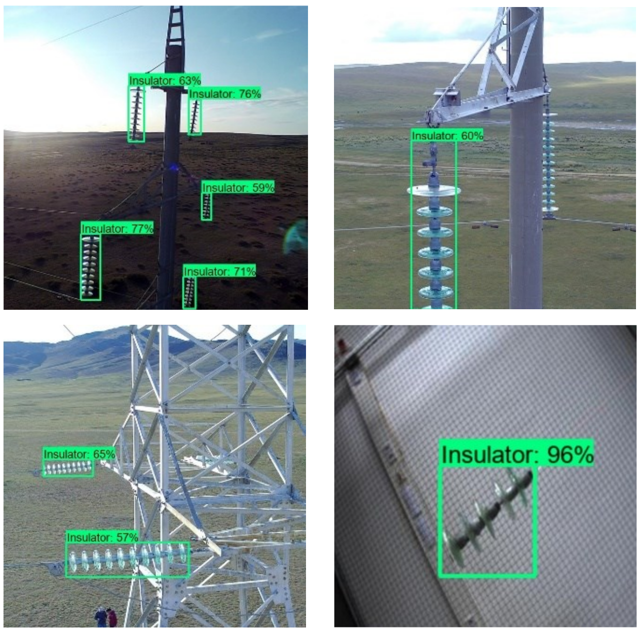


Fig. 3 Power Line Insulator Detection using SSD640

TABLE III - PERFORMANCE EVALUATION ON NVIDIA JETSON NANO

Model		Power Mode	FPS	Avg. Power (W)
SSD300	**Passive DL**	5W	4.2	3.80
	Iteration 1		5.6	3.06
	Iteration 8		5.2	3.50
	Passive DL	10W	8.6	6.81
	Iteration 1		9.0	6.72
	Iteration 8		8.7	6.46
SSD640	**Passive DL**	5W	2.2	3.80
	Iteration 1		2.7	3.58
	Iteration 8		2.6	3.37
	Passive DL	10W	3.4	7.23
	Iteration 1		4.1	7.91
	Iteration 8		3.9	7.77

10W power mode. These modes correspond to an average power consumption of 3.8W and 7.23W, respectively. On the other hand, the active DL approach, even on the model from the 8th iteration, demonstrates a comparable inference rate of 2.6 fps on the 5W power mode and 3.9 fps on the 10W power mode. The power utilization remains relatively efficient, amounting to 3.37W and 7.77W for the respective power modes.

V. CONCLUSION

This research introduced an approach for automated insulator detection utilizing both active and passive learning methodologies within the framework of the SSD. It demonstrated that active learning could significantly lessen the burdensome task of annotating extensive datasets, particularly in specialized applications where the procurement and labelling of data are both costly and labour-intensive.

The outcomes highlight that both the SSD300 and SSD640 models, when trained via active learning strategies, could produce noteworthy results, even when using a limited number of labelled images. Despite a minor trade-off in the frame rate due to a higher image resolution, the SSD640 model exhibited superior precision and recall. Both models adequately fulfilled real-time requirements when implemented on the NVIDIA Jetson Nano platform.

Future efforts will concentrate on enhancing the active learning selection process, exploring the application of other low-energy edge devices, and assessing the effects of different object detection architectures. In general, this research makes a substantial contribution to the application of active learning in computer vision tasks, setting the stage for more effective use of limited labelled data resources in visual inspection tasks.

References

[1] F. Xia, H. Tie, E. Shi, Q. Zhang, and L. Wang, “Insulator recognition based on mathematical morphology and Bayesian segmentation,” in 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE, Oct. 2017, pp. 1–5. doi: 10.1109/CISP-BMEI.2017.8302063.

[2] Y. Zhai, D. Wang, M. Zhang, J. Wang, and F. Guo, “Fault detection of insulator based on saliency and adaptive morphology,” Multimed Tools Appl, vol. 76, no. 9, pp. 12051–12064, May 2017, doi: 10.1007/s11042-016-3981-2.

[3] X. Li, H. Su, and G. Liu, “Insulator Defect Recognition Based on Global Detection and Local Segmentation,” IEEE Access, vol. 8, pp. 59934–59946, 2020, doi: 10.1109/ACCESS.2020.2982288.

[4] M. Oberweger, A. Wendel, and H. Bischof, “Visual Recognition and Fault Detection for Power Line Insulators,” 2014.

[5] S. Liao and J. An, “A Robust Insulator Detection Algorithm Based on Local Features and Spatial Orders for Aerial Images,” IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 5, pp. 963–967, May 2015, doi: 10.1109/LGRS.2014.2369525.

[6] U. Tudevdagva, B. Battseren, W. Hardt, and G. V. Troshina, “Image Processing Based Insulator Fault Detection Method,” in 2018 XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE), IEEE, Oct. 2018, pp. 579–583. doi: 10.1109/APEIE.2018.8545429.

[7] R. Kundu, S. Das, and S. Kumar Mishra, “A Comprehensive Review of Single and Double Stage Object Detection Algorithms in Self Driving Cars,” in 2022 IEEE 7th International conference for Convergence in Technology (I2CT), IEEE, Apr. 2022, pp. 1–8. doi: 10.1109/I2CT54291.2022.9824384.

[8] W. Zhao, M. Xu, X. Cheng, and Z. Zhao, “An Insulator in Transmission Lines Recognition and Fault Detection Model Based on Improved Faster RCNN,” IEEE Trans Instrum Meas, vol. 70, pp. 1–8, 2021, doi: 10.1109/TIM.2021.3112227.

[9] S. Li, H. Zhou, G. Wang, X. Zhu, L. Kong, and Z. Hu, “Cracked Insulator Detection Based on R-FCN,” J Phys Conf Ser, vol. 1069, p. 012147, Aug. 2018, doi: 10.1088/1742-6596/1069/1/012147.

[10] G. Han et al., “Improved Algorithm for Insulator and Its Defect Detection Based on YOLOX,” Sensors, vol. 22, no. 16, p. 6186, Aug. 2022, doi: 10.3390/s22166186.

[11] X. Liu, Y. Li, F. Shuang, F. Gao, X. Zhou, and X. Chen, “ISSD: Improved SSD for Insulator and Spacer Online Detection Based on UAV System,” Sensors, vol. 20, no. 23, p. 6961, Dec. 2020, doi: 10.3390/s20236961.

[12] Batbayar Battseren, Mohamed Salim Harras, and Soaibuzzman, “Deep-Learning-Based Insulator Detector for Edge Computing Platforms,” in International Symposium on Computer Science, Computer Engineering and Educational Technology (ISCSET-2021), Chemnitz: TUDpress, Verlag der Wissenschaften GmbH, 2021, pp. 21–28. Accessed: Jun. 20, 2023. [Online]. Available: https://www.researchgate.net/publication/357877282_Deep-Learning-Based_Insulator_Detector_for_Edge_Computing_Platforms

[13] S. Saleh, H. Saleh, M. Nazari, and W. Hardt, “Outdoor Navigation for Visually Impaired based on Deep Learning,” Jul. 2019.

[14] S. Saleh, S. Khwandah, A. Heller, W. Hardt, and A. Mumtaz, “Traffic Signs Recognition and Distance Estimation using a Monocular Camera,” Jul. 2019.

[15] W. Liu et al., “SSD: Single Shot MultiBox Detector,” 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.

[16] V. N. Kemajou, A. Bao, and O. Germain, “Wellbore Schematics to Structured Data Using Artificial Intelligence Tools,” in Day 1 Mon, May 06, 2019, OTC, Apr. 2019. doi: 10.4043/29490-MS.

[17] L. Zhang and K. Ma, “A Good Data Augmentation Policy is not All You Need: A Multi-Task Learning Perspective,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 5, pp. 2190–2201, May 2023, doi: 10.1109/TCSVT.2022.3219339.