Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi
Niğde Ömer Halisdemir University Journal of Engineering Sciences

Araştırma makalesi / Research article

www.dergipark.org.tr/tr/pub/ngumuh / www.dergipark.org.tr/en/pub/ngumuh

Power grid inspection using YOLOv8 and advanced machine learning techniques for enhanced defect identification

YOLOv8 ve gelişmiş makine öğrenmesi teknikleri ile elektrik şebekesi denetimi ve hassas hata tespiti

Hatice Okumuş[1,*]

[1] *Karadeniz Technical University, Electrical-Electronics Engineering Department, 61080, Trabzon Türkiye*

Abstract

This study proposes an automated grid inspection method by replacing manual inspections, reducing labor, time and improving accuracy. The YOLOv8 classification model's pooling layer outputs are utilized as a feature extractor. Machine learning techniques particularly k-Nearest Neighbor (k-NN), Random Forest (RF) and Neural Network (NN) are then used to assess these collected features. The methodology is tested on a comprehensive image dataset from Brazil's electrical distribution network. This dataset includes images of insulators set against a variety of backgrounds presenting a realistic challenge in identifying defects in grid structures. The aim is to classify these structures as either defective or normal, a crucial step in maintaining the integrity and safety of power grids. The findings of this study are particularly noteworthy demonstrating that the integration of YOLOv8's feature extraction with NN algorithms significantly elevates the accuracy of defect identification in grid structures. Achieving an accuracy of 97.50% and a sensitivity of 95.83% these results not only validate the efficacy of the proposed method but also underscore its superiority over existing methodologies in the literature using the same dataset. The proposed method has promising potential to enhance automated grid inspection processes and contribute meaningfully to ongoing developments in the field.

Keywords: Grid inspection, YOLOv8 network, Feature extraction, Machine learning, Image processing

Öz

Bu çalışma manuel incelemelerin yerine otomatik bir şebeke denetim yöntemi önererek iş gücünü ve zamanı azaltmayı aynı zamanda doğruluğu artırmayı amaçlamaktadır. YOLOv8 sınıflandırma modelinin havuzlama (pooling) katmanından elde edilen çıktılar özellik çıkarıcı olarak kullanılmıştır. Elde edilen bu özellikler başta k-En Yakın Komşu (k-NN), Rastgele Orman (RF) ve Yapay Sinir Ağı (NN) olmak üzere makine öğrenmesi yöntemleriyle değerlendirilmiştir. Önerilen yöntem, Brezilya elektrik dağıtım şebekesinden elde edilen kapsamlı bir görüntü veri seti üzerinde test edilmiştir. Bu veri seti, farklı arka plan koşullarında çekilmiş izolatör görüntülerini içermekte olup şebeke yapılarındaki arızaların tespiti için gerçekçi zorluklar sunmaktadır. Çalışmanın amacı, bu yapıların arızalı veya normal olarak sınıflandırılmasıdır; bu da enerji şebekelerinin bütünlüğünün ve güvenliğinin korunması açısından kritik bir adımdır. Elde edilen bulgular YOLOv8'in özellik çıkarımı ile NN algoritmalarının entegrasyonunun arıza tespit doğruluğunu önemli ölçüde artırdığını göstermektedir. Yöntem %97.50 doğruluk ve %95.83 duyarlılık değerlerine ulaşarak, aynı veri setini kullanan literatürdeki mevcut yöntemlere kıyasla üstün performans sergilemiştir. Bu sonuçlar, önerilen yöntemin etkinliğini doğrulamakta ve otomatik şebeke denetim süreçlerinin geliştirilmesine önemli katkı sunma potansiyelini ortaya koymaktadır.

Anahtar kelimeler: Elektrik şebekesi denetimi, YOLOv8 ağı, Özellik çıkarma, Makine öğrenmesi, Görüntü işleme

1 Introduction

Insulators are vital components in electrical systems tasked with the crucial role of preventing unwanted flow of current to the earth from its intended path along the conductor. However, with time, mechanical loads, material deterioration and environmental conditions may cause these insulators to acquire flaws [1]. These flaws have the potential to cause catastrophic failures and hazardous circumstances like electrocution, explosions or flames. Additionally, when insulators malfunction, the steady flow of electricity may be disrupted resulting in power interruptions or outages. In addition to being inconvenient, this disruption may have a significant impact on homes, businesses and vital services like emergency response systems and hospitals.

Power inspections play a crucial role in maintaining a safe, efficient and reliable power system. This involves checking grid components such as transformers, cables, towers and most importantly insulators for signs of wear, damage or failure. Inspectors use a variety of tools and techniques including visual inspections, thermal imaging and advanced diagnostic technologies to detect issues like overheating, corrosion or structural weaknesses.

* Sorumlu yazar / Corresponding author, e-posta / e-mail: haticeokumus@ktu.edu.tr
Geliş / Recieved: 26.03.2025 Kabul / Accepted: 23.09.2025 Yayımlanma / Published: 15.01.2026
doi: 10.28948/ngumuh.1666320

The use of drones and machine learning for inspections have enhanced the accuracy and efficiency of defect detection. These technology advancements enable more exact defect classification and detection through more accurate and efficient inspections.

In today's landscape of computer vision and artificial intelligence, the utilization of deep learning models for object detection and classification tasks has become a crucial component of image processing and AI applications [2-5]. This development presents significant opportunities for the inspection of power systems. Deep learning models can be utilized to detect anomalies, damages or maintenance needs from high resolution images of power system infrastructure components. This technological approach allows for more effective monitoring and maintenance of power systems, which helps in preventing outages and failures. It also contributes to both cost reduction and an increase in system reliability.

YOLO (You Only Look Once) represents a significant breakthrough in the field of computer vision, known for its ability to perform object detection [6]. Over the years it has evolved through several versions, each improving in terms of speed and accuracy. These improvements have made YOLO a highly effective and accessible tool for a wide range of applications. To date, no research has employed the YOLOv8 classification model as a feature extractor for detecting defects in insulators. This study, therefore pioneers the use of the YOLOv8 model in this capacity leveraging outputs from its pooling layer to identify insulator defects. Consequently, the primary contributions of this study are as follows:

1. This study is the first to apply the YOLOv8 classification model as a feature extractor for the automated inspection of insulators within electrical distribution networks.
2. Experiments are conducted using the YOLOv8s model pooling layer output, which is then analyzed by k-Nearest Neighbor (k-NN), Random Forest (RF) and Neural Network (NN) machine learning techniques to classify insulators as either defective or normal. Results for Yolov8n, Yolov8s, Yolov8m, Yolov8l and Yolov8x models are also investigated for a comparison.
3. The methodology is tested on a comprehensive image dataset from Brazil's electrical distribution network, achieving a classification accuracy of 97.50% and a sensitivity of 95.83% using Yolov8s along with NN. These results surpassed the performance of existing methods in the literature using the same dataset.

2 Related work

Many studies have been conducted in the literature regarding the use of YOLO in power system inspections [7-9]. Chen and Miao have introduced a technique for post-disaster evaluation of power distribution networks, utilizing UAVs to detect and count poles in video footage. This method applies the YOLO algorithm and a convolutional neural network, with adjustments made to YOLO v3's anchor values and the designation of specific Regions Of Interest (ROI) for drone inspections. It features a counting algorithm that monitors shifts in pole positions across sequential video frames, leading to precise and accurate pole counts, with a detection accuracy exceeding 0.9 [10]. Lu et al. have developed a novel defect detection system for transmission lines aimed at reducing the bandwidth usage and response time delays typically encountered in cloud server-based systems. Their approach is based on an "edge-cloud-end" collaborative model, wherein image detection tasks are transferred to edge devices, thereby boosting the efficiency of data transmission and the overall speed of system responses. The system utilizes a YOLO inspection algorithm, which integrates GhostNetV2 for refining the YOLOv5 model. This setup is further supported by a feature fusion network, which includes dynamic weight allocation and cross-scale connectivity. Implemented on the NVIDIA Jetson Xavier NX platform and optimized with TensorRT, the algorithm excels in accurately detecting defects, even under difficult environmental conditions such as low lighting or snow. The system has achieved a notable mean average precision (mAP) of 94.3% [11].

While YOLO's application in power system inspections is broad a notable subset of research specifically targets insulator defect detection. Wang et al. have proposed an insulator defect detection algorithm based on the YOLOv4 model. The algorithm integrates MobileNet-v2 as its backbone supplemented with a Convolutional Block Attention Module (CBAM) and an Ultra-Lightweight Subspace Attention Module (ULSAM) to improve detection of small or dense targets. Extensive datasets of insulators and defects were collected from the Internet to train the model. The improved YOLOv4 algorithm demonstrates high accuracy and speed in experiments, achieving a mean average precision of 87.48% and an insulator defect recall rate of 79.84% [12]. Souza et al. have introduced a hybrid approach utilizing a modified YOLO algorithm integrated with a ResNet-18 classifier for inspecting power systems. The method was specifically designed to analyze real images of defective power components captured by unmanned aerial vehicles (UAVs). The study utilizes a dataset comprising 1593 images of power grid inspections for supervised training. The hybrid YOLO model, based on YOLOv5x demonstrated superior performance in object detection with a mAP of 0.99262 outperforming other versions like YOLOv5n, YOLOv5s, YOLOv5m and YOLOv5l. Furthermore, in multi classification tasks, the proposed Hybrid-YOLO model has achieved an F1_score of 0.96216 surpassing various other architectures including multiple VGG and ResNet models as well as DenseNet versions and YOLO versions up to YOLOv7 [13]. Another YOLO based method for detecting insulators from UAV-captured aerial images was introduced by Liu et al. [14]. The YOLOv3-dense method incorporates Dense-Blocks and multiscale feature fusion for improved accuracy in varying insulator sizes and complex backgrounds. The YOLOv3-dense network outperforms traditional YOLOv3 and YOLOv2, achieving an impressive average precision of 94.47%. Stefenon et al. have presened YOLOu-Quasi-ProtoPNet, a hybrid method combining YOLOv5 and an optimized Quasi-ProtoPNet for detecting and classifying failed insulators.

Various backbones including VGG-16, VGG-19, ResNet-34, ResNet-152, DenseNet-121 and DenseNet-161 were evaluated for optimizing the Quasi-ProtoPNet structure. The model, particular4y based on DenseNet-161, achieved an F1-score of 0.95165 surpassing other similar models [15]. In the study conducted by Chen et al., Insu-YOLO, an advanced object detection model based on the YOLOv8 network, was specifically designed for insulator defect identification. The model incorporates the GSConv module to reduce computational complexity and a lightweight CARAFE structure for efficient feature up sampling in the neck network. Insu-YOLO has achieved a mean average precision of 95.9%, surpassing the YOLOv8n baseline by 3.95% [16]. Yi et al. proposed a new insulator and defect detection model YOLO-S (YOLO-Small) which addresses low accuracy and high computational demands of previous methods. It features lightweight GSConv, GSbottleneck, and the VoV-GSCSP module for efficiency, and the MaECA attention module for better target detection. The model also uses the SIoU loss function, and replaces the SILU function in YOLOv5s with Mish for improved image integrity. YOLO-S surpasses YOLOv5s with a 4.2% higher mean average precision, 2.1% better accuracy, and a 6.0% reduction in computational load [17]. Wang et al. have presented an enhanced YOLOv4 algorithm for insulator defect detection incorporating new data augmentation and redesigned anchor boxes using the K-means algorithm. This improvement leads to a 37.2% increase in detection precision. The algorithm outshines existing methods like SSD, Faster-RCNN and previous YOLO versions achieving 99.08% mean average precision and 56 fps. Its robustness across different lighting and environmental conditions makes it highly suitable for industrial applications [18]. In another study Zhang et al have proposed an improved YOLOv8s-based algorithm for insulator defect detection. This enhancement involves integrating a Multi-Scale Large-Kernel Attention (MLKA) module for better focus on various scale features and low-level maps employing lightweight GSConv convolution along with the GSC_C2f module to reduce computational complexity and memory usage. Additionally, an enhanced SIoU loss function is utilized to optimize the detection performance. The improved model has excelled in drone aerial photography for insulator defect detection, achieving a mAP of 99.22% and 55.73 frames per second (FPS). This represents a significant improvement over the original YOLOv8s and YOLOv5s models with increases of 2.18% and 2.91% in mAP respectively [19]. Chang et al.have developed a technique that integrates YOLOv7 with a multi-UAV system to enhance the identification of faults in transmission lines. They evaluated this method using a diverse dataset, which includes images from web crawlers, public sources and direct captures focusing on three common types of transmission line defects which are cracked insulators, insulator self-blasts and bird nests. In comparison to the YOLOv5-S model, their YOLOv7 variant demonstrates improved performance with a 1.2% increase in accuracy, a 4.3% rise in recall and a 4.1% enhancement in mAP. Notably, the method achieves a robust mAP of 0.886 showing high effectiveness even in challenging imaging conditions with complex backgrounds and noise [20].

Despite the substantial contributions of these studies to defect detection there appears to be a noticeable gap in the literature regarding the detailed classification of defective structures, particularly those involving insulators. Addressing this gap, this study proposes a YOLOv8 approach for the classification of defective network structures. The YOLOv8 classification model is not directly used; instead it is employed as a tool for feature extraction specifically leveraging the outputs from its pooling layer for more in-depth analysis. These extracted features are then processed using advanced machine learning techniques, specifically k-NN, RF and NN to distinguish between defective and normal grid structures.

The methodology is tested on a comprehensive image dataset obtained from Brazil's electrical distribution network. This dataset, comprising images of insulators against diverse backgrounds, presents a realistic scenario for testing defect identification capabilities in grid structures. The primary objective is to enhance the classification accuracy of these structures as either defective or normal, a crucial aspect in upholding the integrity and safety of power grids. Preliminary results indicate that integrating YOLOv8's feature extraction capabilities especially with NN substantially improves the accuracy in defect identification, underscoring the potential of this approach to revolutionize automated grid inspection processes.

The organization of the remaining sections of the paper is structured in the following manner: Section 3 provides an in-depth exploration of the proposed methodology, detailing the YOLOv8 models and the classification techniques employed. In Section 4, there is a comprehensive assessment of the system's performance, where the various performance metrics are explained and analyzed. The paper ends in Section 5, which offers a conclusion, summarizing the principal discoveries and insights derived from the research.

3 Proposed methodology

The proposed methodology for the classification of power grid structures is visually represented in Figure 1. The cornerstone of our methodology is the use of YOLOv8 classification model as a feature extractor. This model, known for its efficiency in object detection tasks, has been adapted in our study to extract pertinent features from images of power grid structures. The unique aspect of our approach lies in utilizing the outputs from the pooling layer of YOLOv8 which are critical in capturing the essential characteristics of the images. Once these features are extracted, they are subjected to further analysis using well established machine learning algorithms such as k-NN, RF and NN. Computational tasks are carried out on a system with 8 GB of RAM and a 3.40 GHz Intel Core i7 processor providing strong processing power and effective analysis capabilities. Detailed explanations of each stage of this methodology are provided in subsequent subsections.

3.1 You only look once version 8 (YOLOv8)

The You Only Look Once (YOLO) algorithm proposed by Redmon et al. in 2016 [6] marked a significant departure

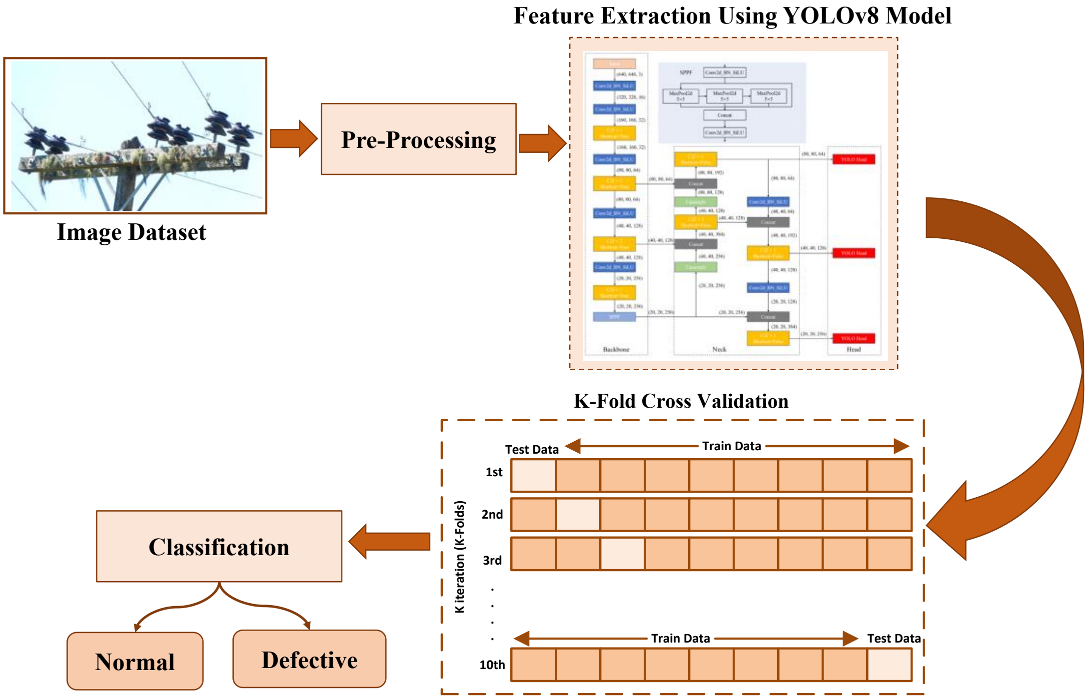


Figure 1. General block diagram of the proposed approach

from conventional object detection techniques. Unlike traditional object detection methods which involve multiple scans or separate stages for detection and classification, YOLO accomplishes this in a single forward pass of a neural network. This integration not only accelerates the detection process but also strikes an effective balance between speed and accuracy, making YOLO highly suitable for real-time applications such as video surveillance and autonomous vehicles.

As the algorithm has evolved, various iterations have been released, each enhancing its capabilities [21]. The most recent iteration, YOLOv8, developed by Ultralytics in 2023, represents a notable advancement. YOLOv8 enables detection, segmentation, pose estimation, tracking and classification further augmenting its versatility and utility in diverse applications [22].

YOLOv8's architecture aligning with the detection principles of its predecessors YOLOv5 [23] and YOLOv7 [24] is depicted in Figure 2. This model is structured around four principal elements: the input, backbone, neck and head. Initially, the input segment undertakes the task of pre-processing the image through various procedures including data enhancement. This pre-processed image is then relayed to the backbone, which is responsible for extracting features. Following this, the neck module amalgamates these extracted features to generate feature representations in multiple sizes (such as large, medium, and small). In the final stage, these amalgamated features are forwarded to the detection head, culminating in the output of the detection results [16]. In terms of model offerings, YOLOv8 presents a suite of five primary models: YOLOv8n (Nano), YOLOv8s (Small), YOLOv8m (Medium), YOLOv8l (Large) and YOLOv8x (Extra Large). Amongst these YOLOv8n stands out as the fastest and smallest model, whereas YOLOv8x as the slowest but most accurate. For training purposes, detection and segmentation models are nurtured using the COCO dataset with images of 640 pixels in size, whereas the ImageNet dataset, featuring an image input size of 224, is employed for the classification models.

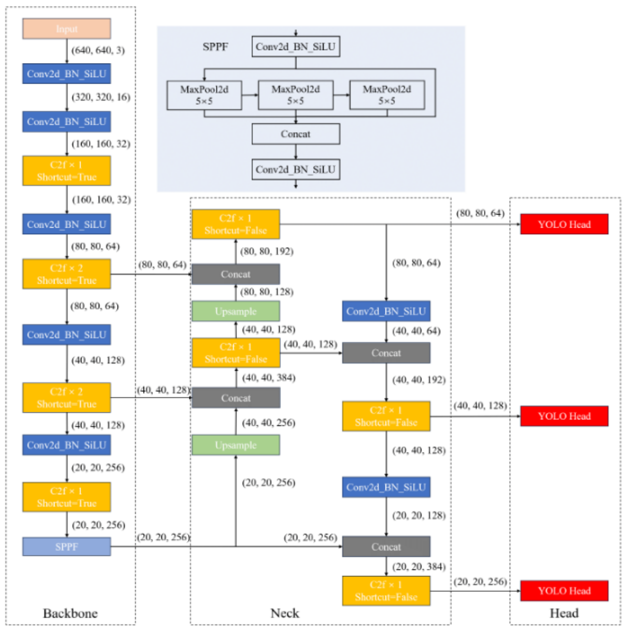


Figure 2. YOLOv8 structure [16]

In this research, the YOLOv8s classification model is utilized for feature extraction purposes.

3.2 Using YOLOv8 as a feature extractor

As stated before in this study, we delve into the novel use of YOLOv8, primarily recognized for its object detection prowess, to extract features for image classification. Our key innovation involves utilizing the outputs from the model's pooling layer, instead of its final outputs. This approach aims to capture a deeper and more representative set of features from the images.

During data processing, the input images are resized to 640×640 pixels and normalized to scale pixel values between 0 and 1. Although YOLOv8 classification models are typically pre-trained on 224×224 inputs, this resolution was chosen to obtain more detailed spatial features. Feature extraction is performed by capturing the activations from the global average pooling layer located at the transition between the backbone and the classification head, which provides a compact representation of high-level visual features.

The outputs from the pooling layer, are essentially activation maps that reveal vital visual characteristics identified by the model. The dimensions of these maps, influenced by YOLOv8's architecture, typically reduce the spatial size of the input while maintaining depth, which contains essential image information. The pooling layer's extracted features are converted into a one-dimensional format resulting in a 1x1280 feature matrix for each image in this study. This matrix serves as the base for subsequent classification processes.

3.3 Classification methods

To differentiate the structures as either normal or defective, established machine learning classification methods are utilized.

3.3.1 k-nearest-neighbor (k-NN)

The k-NN algorithm is a simple yet effective method he k-NN algorithm is a simple yet effective method used in supervised learning for both classification and regression problems across categorical or continuous datasets. While k-NN is renowned for its simplicity and ease of understanding, it can suffer from slow performance on large datasets and be affected by the curse of dimensionality. Moreover, it may not perform well with improperly normalized data. Despite these limitations, k-NN finds extensive applications ranging from financial forecasting and medical diagnosis to image recognition and recommendation systems [25-27].

k-NN functions by considering the 'k' nearest neighbors to a given data point to formulate a prediction or classification. Selecting the appropriate 'k' value is crucial as it significantly affects the algorithm's effectiveness. Typically, k-NN uses distance metrics such as Euclidean, Manhattan or Minkowski to identify the closest neighbors. The process involves choosing a 'k' value computing the distance between the data point and all other points in the dataset, pinpointing the nearest 'k' neighbors and then for classification, determining the most common class among these neighbors as the prediction. In regression scenarios, it averages their output values for the prediction.

In this study, the Euclidean distance metric is utilized to calculate the distances between neighbors. The formula for computing the Euclidean distance between m-dimensional points a and b is provided in Equation (1).

$$d_E(a,b) = \sqrt{\sum_{i=1}^{m}(a_i - b_i)^2} \tag{1}$$

3.3.2 Random forest (RF)

The Random Forest algorithm is a form of ensemble learning, distinguished for its effectiveness in both classification and regression tasks [28]. It is highly valued for its accuracy and robustness. At the heart of the Random Forest method is 'bagging' also known as bootstrap aggregating. This process entails training each tree in the forest (n trees) on a uniquely drawn sample from the training set, selected randomly with replacement. Further injecting randomness, the algorithm selects a random subset of features (m features) at every decision point or split in the tree. This aspect of randomness is crucial as it significantly enhances the model's resilience to overfitting. Random Forest evaluates the performance of each individual decision tree using its internal errors, commonly known as Out-of-Bag (OOB) errors. The weighting of trees is based on their error rate, with lower-error trees receiving higher weights and higher-error trees assigned lower weights. For predictions, each tree contributes a weighted vote and the average is taken as the final output while in classification scenarios, the class that garners the majority of these weighted votes is chosen as the output.

3.3.3 Neural network (NN)

Neural networks (NNs), mirroring the architecture of the human brain, are a fundamental element in contemporary artificial intelligence. Composed of units or neurons linked together, these networks are structured in layers, including input, hidden, and output layers [29]. Their learning mechanism involves a training phase where the network modifies its weights in response to input data, and backpropagation, which fine-tunes these adjustments to reduce errors. Remarkably proficient at complex functions such as image recognition and language processing, NNs thrive by discerning patterns in large datasets [30-32]. Their versatility and growing complexity point to a future where they might tackle increasingly intricate and human-like challenges. Despite the wide range of neural network architectures available, this paper focuses on the feedforward neural network model as given in Figure 3. They are the most basic form of artificial neural networks, characterized by unidirectional connections between nodes that avoid any cycles. This setup resembles a one-way flow of information, starting at the input layer, possibly passing through various hidden layers, and culminating at the output layer. The non-cyclic nature of these networks renders them a relatively simple and manageable computational model, well-suited for tasks like pattern recognition and data classification.

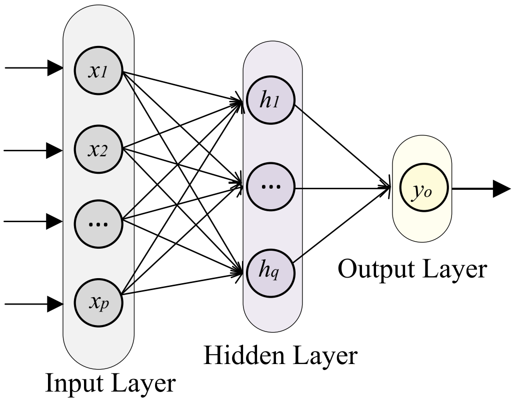


Figure 3. Structure of feedforward neural network model

The decision to employ this type of neural network in our study is driven by its straightforwardness, efficiency and the specific problem we are addressing. The feedforward neural networks uncomplicated architecture facilitates efficient computation while retaining a high level of accuracy in predictive tasks. This makes it particularly suitable for the scope of this research where a balance between complexity and performance is essential.

In this study, the k-NN algorithm was applied with k = 5, while the Random Forest model consisted of 100 decision trees as these configurations were found to offer an effective trade off between accuracy and processing time during preliminary evaluations. For the neural network a feedforward structure was utilized and trained in MATLAB using default parameters with the number of training epochs set to a maximum of 100.

3.4 Data description

In this study a dataset comprising 240 images from Southern Brazil based on reported issues by Santa Catarina's electric utility [33] is analyzed. These images, captured during grid inspections, depict both insulation conditions and pollution levels on the grid structures. The dataset is evenly divided, featuring 120 images of defective structures and 120 of normal ones, providing a comprehensive view of the grid's state. These images provide a valuable resource for assessing the grid's condition and understanding common issues in the region's electric utility infrastructure. Samples of the dataset is given in Figure 4. The first line includes structures with defects, while the second line shows normal structures.

3.5 Performance metrics

The method under discussion, as noted earlier, involves distinguishing between defective and normal structures through classification. To assess the efficacy of the technique used in such classification tasks, a range of metrics are can be employed. A commonly utilized tool is the confusion matrix, which compares the actual target values with those predicted by the model, allowing for an easy assessment of how well the model is performing. The confusion matrix consists of 4 categories for a binary classification problem, which are True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). True Positive (TP) represents the number of positive instances correctly classified as positive. True Negative (TN) is the number of negative instances correctly classified as negative. False Positive (FP) signify the number of negative instances incorrectly classified as positive. False Negative (FN) stands for the number of positive instances incorrectly classified as negative.

The metrics of accuracy (Acc), sensitivity (Sen), specificity (Spec), precision (Prec), and the F1 score (F1) defined in Table 1 can be derived from the confusion matrix to evaluate a classifier's performance.

In this study positive instances indicate images containing defective structures while negative instances indicate images containing normal structures. To improve the accuracy of our performance assessment, we utilized the K-Fold Cross-Validation method. K-Fold Cross-Validation is a method used in machine learning to accurately assess a model's performance. It involves splitting the entire dataset into 'K' parts, or folds. In each of the K iterations, a different fold is used as a validation (test) set to test the model, while the remaining K-1 folds are combined to form the training set. This cycle ensures that each data point is used exactly once for validation and K-1 times for training, which minimizes both bias and variance in the model's evaluation. The overall performance of the model is calculated as the average of the results from all K iterations. For example, if accuracy is the performance measure, the final result is the mean accuracy value obtained across all folds. This method is particularly useful for small datasets, as it maximizes both the training and validation data, offering a thorough and balanced view of the model's effectiveness and its ability to generalize to new data.

Figure 4. Example images from dataset

Table 1. Confusion matrix

Confusion Matrix				*Performance Metrics*	
		Predicted Class		***Acc***	$\frac{TP+TN}{TP+TN+FP+FN}$
		1	*2*	***Sen***	$\frac{TP}{TP+FN}$
Actual Class	*1*	*TP*	*FN*	***Spec***	$\frac{TN}{TN+FP}$
	2	*FP*	*TN*	***Prec***	$\frac{TP}{TP+FP}$
				F1	$\frac{2TP}{2TP+FP+FN}$

4 Results and discussion

Applying the 10-Fold Cross-Validation approach, as described in the previous section, the performance metric results for different machine learning methods presented in Table 2 reveals that employing the YOLOv8s classification model as a feature extractor consistently yields performance levels exceeding 81%. Notably, the k-NN method registers the lowest accuracy at 85.83%, whereas the NN method achieves the highest accuracy reaching 97.50%.

Table 2. Results for each method with YOLOv8s

Metric	Method		
	k-NN	RF	NN
Acc	85.83	87.92	**97.50**
Sen	90.00	88.33	95.83
Spec	81.67	87.50	99.17
Prec	83.08	87.60	99.14
F1	86.40	87.97	97.46

Furthermore, it is apparent that the NN method also attains the highest all the performance metrics.

A comparative analysis is performed among YOLOv8 model variants—Yolov8n, Yolov8s, Yolov8m, Yolov8l, and Yolov8x in order to see the performance differences. The results obtained with k-NN classifier is highlighted in Figure 5. All models showcase an accuracy above 76%, with Yolov8n and Yolov8s slightly ahead at 84.17% and 85.83% respectively. This suggests that the simpler models, Yolov8n and Yolov8s, perform comparably to their more complex counterparts in our dataset. Yolov8s leads in sensitivity, while Yolov8n excels in specificity, indicating better identification of negative cases, though with slightly lower sensitivity than Yolov8s. In precision, Yolov8n outperforms the rest, implying fewer false positives, crucial in certain applications. Additionally, all models demonstrate competitive F1 Scores, with Yolov8s having a marginal lead, making it an ideal choice for scenarios valuing both precision and sensitivity.

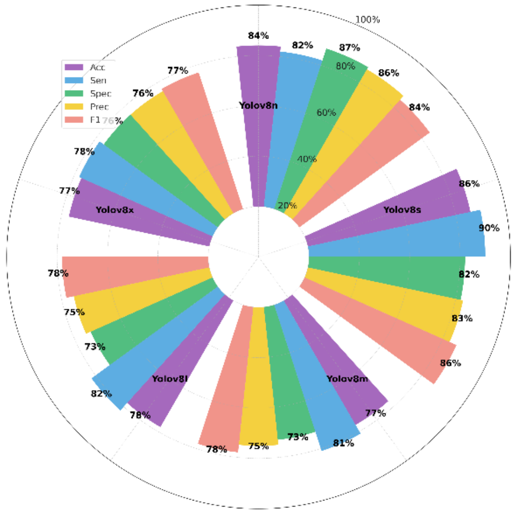


Figure 5. Performance comparison for *k*-NN

Figure 6 presents the performance results using the RF method. Yolov8s achieves the highest accuracy at 87.92%, with Yolov8m at the lower end with 81.67%. Yolov8s tops in sensitivity at 88.33%, showcasing its ability to accurately identify true positives. Yolov8m has the lowest specificity at 84.17%, while Yolov8n leads in precision with 87.61%, indicating its effective true positive identification. Yolov8m, also has the lowest precision at 83.33%. Yolov8s maintains higher F1 Score at approximately 88%, indicating a balanced interplay between precision and sensitivity.

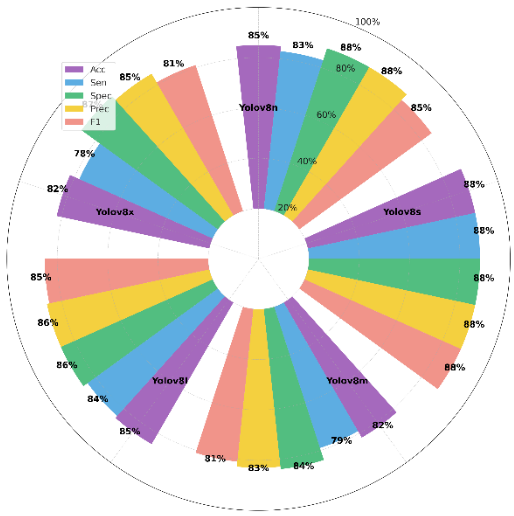


Figure 6. Performance comparison for RF

Figure 7 provides insights into using NN classification for defective structure detection. With an outstanding accuracy peak of over 95%, Yolov8s and Yolov8l stand out in this case, exhibiting constant precision across the dataset. Along with Yolov8l, Yolov8s earns a high sensitivity score of 95.83%, and it also demonstrates its strength with the greatest specificity of 99.17%. Furthermore, Yolov8s is the most accurate in positive predictions with a precision of 99.14%. Yolov8s has an F1 Score nearly 98% indicating its well-rounded performance. Despite lagging in accuracy, sensitivity and F1 Score, Yolov8x still demonstrates significant specificity and precision values over 94% highlighting its effectiveness.

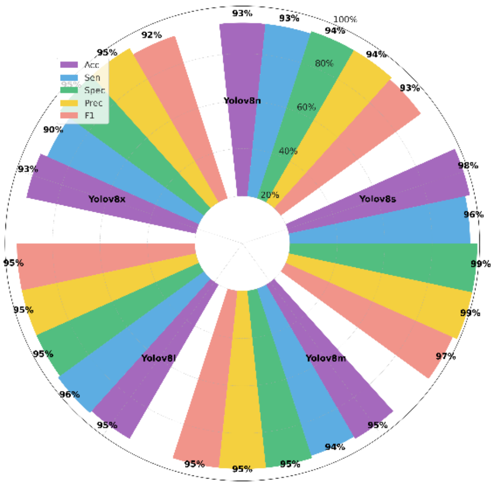


Figure 7. Performance comparison for NN

In summary, Table 3 provides a comprehensive overview detailing both the average performance metrics and the in-depth results for each method. The analysis of table reveals that the k-NN method yields an Acc rate of 80.25%, with its strongest performance in Sen reaching 82.33%. Conversely, the RF method outperforms k-NN demonstrating a higher effectiveness particularly in Spec where it achieves a rate of 86.50%. The average scores across the methods suggest a clear performance hierarchy with the NN method outperforming the k-NN and RF in every metric. Especially in terms of Prec the NN method stands out with an average of 95.59%.

The confusion matrices in Figure 8 provides a detailed comparison of the classification performance of various YOLOv8 models evaluated using three different methods: k-NN, RF and NN. Each subfigure (a, b, and c) corresponds to these methods, applied across YOLOv8 variants (n, s, m, l, and x). For this analysis, the dataset consisted of 240 samples, and feature extraction yielded a final data dimension of 240x1280. As mentioned before 10 fold cross validation strategy is employed to enhance the robustness and reliability of the classification results. The values presented in the confusion matrices reflect the averaged performance metrics obtained from the 10 fold cross validation highlighting the models' generalizability and consistency.

The matrices effectively illustrate the models' ability to differentiate between normal and defective classes. Among the methods again it can be seen that the NN approach consistently demonstrates superior performance as evidences by the higher values along the diagonal of the matrices indicating a more accurate classification with fewer misclassifications when compared to the k-NN and RF methods. Notably, the combination of the YOLOv8s model with the NN method proves to be particularly effective achieving balanced and reliable results across both normal and defective classes.

A comparison has been carried out with the studies in the literature using the same dataset in order to classify defective power system structures. The results of these comparative studies are detailed in Table 4.

Table 4. Comparison with the studies in the literature

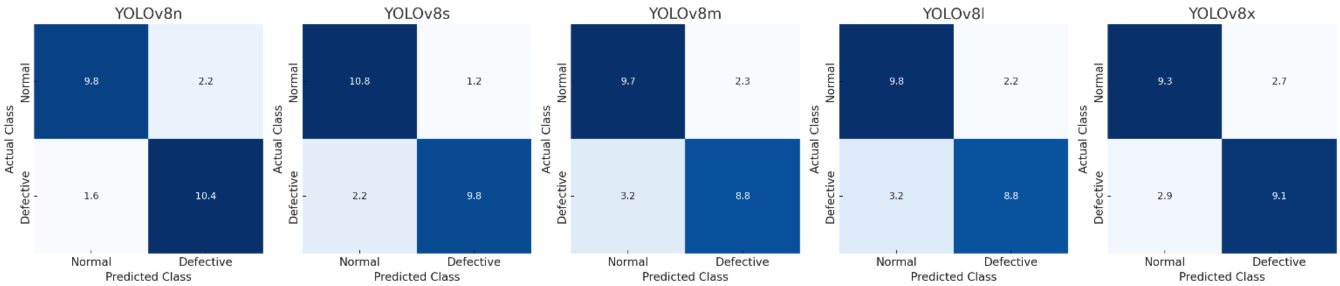
Reference	Performance Metrics using NN				
	Acc	Sen	Spec	Prec	F1
[33]	84.72	83.33	-	85.71	84.50
[34]	97.22	99.99	-	94.73	97.29
Proposed	**97.50**	95.83	**99.17**	**99.14**	**97.46**

The study cited in [33] employed an Inception v3 model, using a where augmented dataset where the augmentation was achieved through segmentation and edge detection techniques. Meanwhile, the research in [34] implemented a SemiProtoPNet model for structure classification, using a VGG-19 baseline model with a downsized dataset of 160 samples.

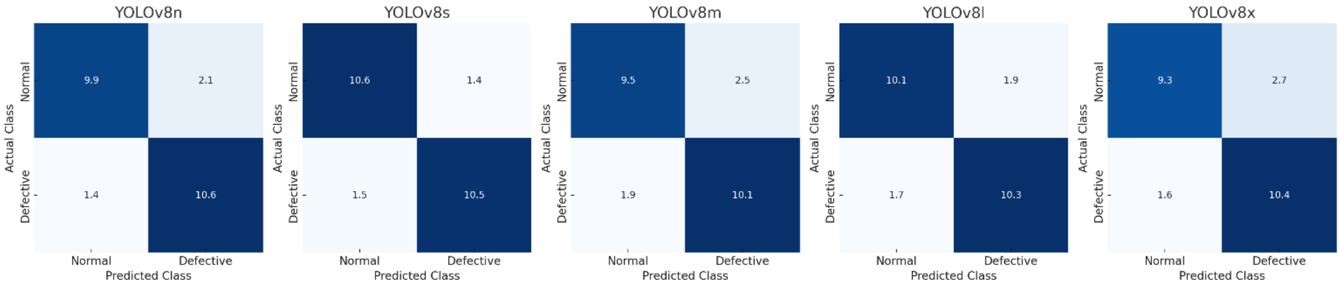
Upon examining the outcomes, it is noted that the approach in [33] achieved a noteworthy accuracy of 84.72% and a precision of 85.71%. Although these results are commendable, they highlight potential areas for improvement, especially in terms of sensitivity and the F1 score. In comparison, the method presented in [34] excelled in sensitivity, reaching an almost perfect 99.99%, which indicates an exceptional rate of correctly identifying true positives. However, this method's accuracy and F1 score, though high, fell short of this level of excellence, pointing to

Table 3. Results for each method using different YOLOv8 models

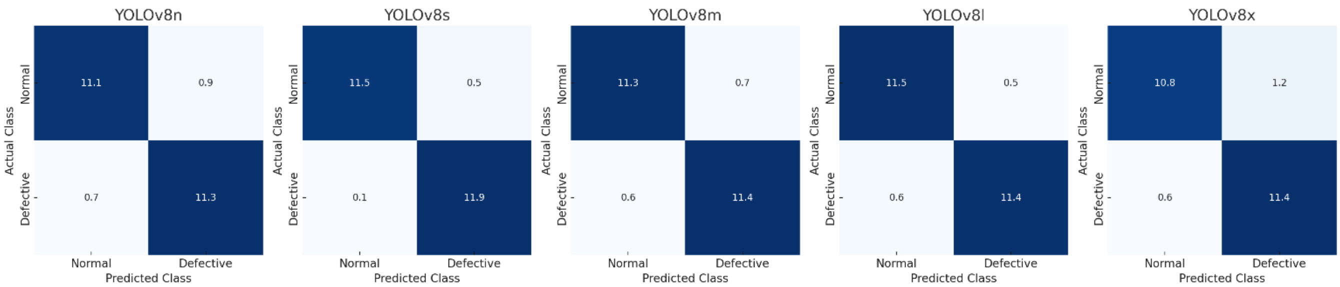
Feature Extractor Model	Method														
	k-NN					RF					NN				
	Acc	Sen	Spec	Prec	F1	Acc	Sen	Spec	Prec	F1	Acc	Sen	Spec	Prec	F1
Yolov8n	84.17	81.67	86.67	85.96	83.76	85.42	82.50	88.33	87.61	84.98	93.33	92.50	94.17	94.07	93.28
Yolov8s	85.83	90.00	81.67	83.08	86.40	87.92	88.33	87.50	87.60	87.97	**97.50**	**95.83**	**99.17**	**99.14**	**97.46**
Yolov8m	77.08	80.83	73.33	75.19	77.91	81.67	79.17	84.17	83.33	81.20	94.58	94.17	95.00	94.96	94.56
Yolov8l	77.50	81.67	73.33	75.38	78.40	85.00	84.17	85.83	85.59	84.87	95.42	95.83	95.00	95.04	95.44
Yolov8x	76.67	77.50	75.83	76.23	76.86	82.08	77.50	86.67	85.32	81.22	92.50	90.00	95.00	94.74	92.31
Average	80.25	82.33	78.17	79.17	80.67	84.42	82.33	86.50	85.89	84.05	94.67	93.67	95.67	95.59	94.61



a) with the *k*-NN method



b) with the RF method



c) with the NN method

Figure. 8. Average confusion matrices for different YOLOv8 models

a possible compromise between sensitivity and other performance metrics.

Our methodology, in contrast, demonstrates an excellent performance with a higher accuracy of 97.50%, a lower sensitivity of 95.83%, and an impressive F1 score of 97.46%. Additionally, our specificity rate is at 99.17% suggesting a strong ability to identify true negatives, although this metric was not directly comparable with the other studies due to lack of data.

Overall, the suggested approach performs better than the cited research in the majority of reported measures, particularly in specificity and precision, demonstrating its efficacy and dependability. The lack of specificity information in references [33] and [34] may indicate that the range of their stated performance is possibly constrained.

The proposed approach in this paper is especially beneficial in practical scenarios, where both false positives and negatives can have serious implications. The analysis underscores the effectiveness of our proposed method showing not only its competitiveness with existing approaches but also its consistent performance across multiple metrics. However, it is important to note that no detailed analysis regarding computational load or processing speed was conducted in this study. Since YOLOv8 was used solely for feature extraction rather than as a full end-to-end model, we expect that the overall complexity and resource requirements are lower than typical deep learning frameworks employing full detection pipelines.

The method can be integrated into existing power grid inspection workflows with minimal disruption. In particular, in systems that utilize UAVs for image collection, the proposed feature extraction and classification steps can operate independently as a separate module. While promising, the method still faces practical limitations in real-world deployments especially due to variable image quality resulting from lighting conditions, weather, and camera angles. To mitigate these challenges, future work should focus on incorporating adaptive preprocessing techniques, robust data augmentation, and testing scenarios under diverse environmental conditions to assess generalizability and resilience.

5 Conclusions

This research primarily focuses on enhancing the efficiency and efficacy of power grid inspections by transitioning from traditional, manual site evaluations to a technologically advanced approach. YOLOv8, representing the most advanced iteration in the YOLO series, is employed in this study not for its conventional image detection usage, but specifically for the classification of structural defects. In this application YOLOv8 is utilized as a feature extractor where the outputs from its pooling layers serve as the primary features. Subsequently, these extracted features are analysed using traditional classification methods such as k-NN, RF and NN. This approach provides a distinctive combination by using YOLOv8 to extract important features and then classifying them using well-established machine learning algorithms.

Our methodology was rigorously tested using a diverse dataset of images from Brazil's electrical networks which presented varied backgrounds and complexities inherent in identifying grid anomalies. The primary objective was to discern operational from defective components within the grid, a critical factor in maintaining the safety and reliability of power supply systems.

The results of our study are promising, particularly it has been observed that integrating the feature matrix obtained from the pooling layer of the YOLOv8s model with NN yields superior results with 97.50% Acc and 95.83% Sen compared to other models and methods. This finding emphasizes the potential of the usage of YOLOv8 models as feature extractors in the field of image classification, especially for specific applications such as in this study. Moreover, these results can guide researchers working on similar problems and lay a foundation for future studies that may involve a more detailed examination of the YOLOv8 model.

We recognize that the comparatively limited dataset (240 images total, balanced between faulty and normal classes) is a limitation of this work. Even though 10-fold cross-validation was used to lessen overfitting and produce a more accurate performance estimate, the results' generalizability might be hampered by the small dataset size. However, it's crucial to remember that the dataset is made up of genuine photos taken from Brazil's electrical system which gives the results more legitimacy and usefulness. In order to further evaluate the reliability and scalability of the suggested approach it would be advantageous to work with bigger and more varied datasets in future studies encompassing other insulator kinds, ambient conditions and imaging setups.

Conflict of interest

The authors declare no competing interests.

Similarity ratio (iThenticate): %14

References

[1] X. B. Huang, Y. H. Gao, Y. Zhang, L. Zhao, Y. Wu, et al., Glass insulator target recognition algorithm based on joint component grayscale algorithm and deep learning. Electrical Power Automation Equipment, 42(4), 203-209, 2022. https://doi.org/10.16081/j.epae.202201032.

[2] P. J. Seegehalli, B. N. Krupa, Deep hybrid architectures and DenseNet35 in speaker-dependent visual speech recognition. Signal, Image and Video Processing, 18(5), 4875-4889, 2024. https://doi.org/10.1007/s11760-024-03123-2.

[3] E. Baykal Kablan, S. Ayas, Skin lesion classification from dermoscopy images using ensemble learning of ConvNeXt models. Signal, Image and Video Processing, 1-9, 2024. https://doi.org/10.1007/s11760-024-03321-y.

[4] S. M. Omer, K. Z. Ghafoor, S. K. Askar, Lightweight improved yolov5 model for cucumber leaf disease and pest detection based on deep learning. Signal, Image and Video Processing, 18(2), 1329-1342, 2024. https://doi.org/10.1007/s11760-023-02865-9.

[5] E. Ergün, Artificial Intelligence Approaches for Accurate Assessment of Insulator Cleanliness in High-Voltage Electrical Systems. Electrical Engineering, 1-16, 2024. https://doi.org/10.1007/s00202-024-02691-3.

[6] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779-788, 2016. https://doi.org/10.1109/CVPR.2016.91.

[7] Z. Lin, W. Chen, L. Su, Y. Chen, T. Li, HS-YOLO: Small object detection for power operation scenarios. Applied Sciences, 13(19), 11114, 2023. https://doi.org/10.3390/app131911114.

[8] Y. Yang, S. Geng, C. Cheng, et al., An edge algorithm for assessing the severity of insulator discharges using a lightweight improved YOLOv8. Journal of Electrical Engineering & Technology, 2024. https://doi.org/10.1007/s42835-024-02021-4.

[9] Y. Zhang, Y. Dou, K. Yang, et al., Insulator defect detection based on BaS-YOLOv5. Multimedia Systems, 30, 212, 2024. https://doi.org/10.1007/s00530-024-01413-w.

[10] B. Chen, X. Miao, Distribution line pole detection and counting based on YOLO using UAV inspection line video. Journal of Electrical Engineering & Technology, 15, 441-448, 2020. https://doi.org/10.1007/s42835-019-00230-w.

[11] L. Lu, Z. Chen, R. Wang, L. Liu, H. Chi, Yolo-inspection: Defect detection method for power transmission lines based on enhanced YOLOv5s. Journal of Real-Time Image Processing, 20(5), 104, 2023. https://doi.org/10.1007/s11554-023-01360-1.

[12] S. Wang, H. Xu, M. Shao, L. Zhao, Insulator defect detection based on improved you-only-look-once v4 in complex scenarios. Journal of Electronic Imaging, 32(2), 023029, 2023. https://doi.org/10.1117/1.JEI.32.2.023029.

[13] B. J. Souza, S. F. Stefenon, G. Singh, R. Z. Freire, Hybrid-YOLO for classification of insulators defects in transmission lines based on UAV. International Journal of Electrical Power & Energy Systems, 148, 108982, 2023. https://doi.org/10.1016/j.ijepes.2023.108982.

[14] C. Liu, Y. Wu, J. Liu, Z. Sun, Improved YOLOv3 network for insulator detection in aerial images with diverse background interference. Electronics, 10(7), 771, 2021. https://doi.org/10.3390/electronics10070771.

[15] S. F. Stefenon, G. Singh, B. J. Souza, R. Z. Freire, K.-C. Yow, Optimized hybrid YOLOu-Quasi-ProtoPNet

for insulators classification. IET Generation, Transmission & Distribution, 17, 3501-3511, 2023. https://doi.org/10.1049/gtd2.12886.
[16] Y. Chen, H. Liu, J. Chen, J. Hu, E. Zheng, Insu-YOLO: An insulator defect detection algorithm based on multiscale feature fusion. Electronics, 12(15), 3210, 2023. https://doi.org/10.3390/electronics12153210.
[17] W. Yi, S. Ma, R. Li, Insulator and defect detection model based on improved YOLO-S. IEEE Access, 11, 93215-93226, 2023. https://doi.org/10.1109/ACCESS.2023.3309693.
[18] S. Wang, X. Zou, W. Zhu, L. Zeng, Insulator defects detection for aerial photography of the power grid using You Only Look Once algorithm. Journal of Electrical Engineering & Technology, 1-14, 2023. https://doi.org/10.1007/s42835-023-01385-3.
[19] L. Zhang, B. Li, Y. Cui, Y. Lai, J. Gao, Research on improved YOLOv8 algorithm for insulator defect detection. Journal of Real-Time Image Processing, 21(1), 22, 2024. https://doi.org/10.1007/s11554-023-01401-9.
[20] R. Chang, P. Xiao, H. Wan, S. Li, C. Zhou, F. Li, A transmission line defect detection method based on YOLOv7 and multi-UAV collaboration platform. Journal of Electrical and Computer Engineering, 2023(1), 9943589, 2023. https://doi.org/10.1155/2023/9943589.
[21] M. Hussain, YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines, 11(7), 677, 2023. https://doi.org/10.3390/machines11070677.
[22] G. Jocher, A. Chaurasia, J. Qiu, YOLO by Ultralytics. 2023.
[23] X. Zhu, S. Lyu, X. Wang, Q. Zhao, TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 2778-2788, 2021. https://doi.org/10.1109/ICCVW54120.2021.00312.
[24] C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao, YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7464-7475, 2023. https://doi.org/10.48550/arXiv.2207.02696.
[25] M. Krishna Satya Varma, K. Raja, N. K. Kameswara Rao, Hybrid optimal joint spatial-spectral hyperspectral image classification using modified DHO-based GIF with JRKNN. The Imaging Science Journal, 71(7), 585-598, 2023. https://doi.org/10.1080/13682199.2023.2187515.
[26] M.F. Keskenler, E. Çelik & D. Dal, A New Multi-Layer Machine Learning (MLML) Architecture for Non-invasive Skin Cancer Diagnosis on Dermoscopic Images. J. Electr. Eng. Technol, 19, 2739–2755, 2024. https://doi.org/10.1007/s42835-023-01758-8.
[27] M. A. Araaf & K. Nugroho, Comprehensive analysis and classification of skin diseases based on image texture features using K-nearest neighbors algorithm. Journal of Computing Theories and Applications, 1(1), 31-40, 2023. https://doi.org/10.33633/jcta.v1i1.9185.
[28] F. Waris, F. Da & S. Liu, Deep learning based features extraction for facial gender classification using ensemble of machine learning technique. Multimedia Systems, 30(200), 2024. https://doi.org/10.1007/s00530-024-01399-5.
[29] M. K. Tripathi & D. D. Maktedar, Internal quality assessment of mango fruit: an automated grading system with ensemble classifier. The Imaging Science Journal, 70(4), 253-272, 2022. https://doi.org/10.1080/13682199.2023.2166657.
[30] P. Singh, S. K. Borgohain, A. K. Sarkar, J. Kumar & L. D. Sharma, Feed-forward deep neural network (FFDNN)-based deep features for static malware detection. International Journal of Intelligent Systems, 2023(1), 9544481, 2023. https://doi.org/10.1155/2023/9544481.
[31] S. Avinash, H. N. Naveen Kumar, M. S. Guru Prasad, R. Mohan Naik & G. Parveen, Early Detection of Malignant Tumor in Lungs Using Feed-Forward Neural Network and K-Nearest Neighbor Classifier. SN Computer Science, 4(2), 195, 2023. https://doi.org/10.1007/s42979-022-01606-y.
[32] D. Konar, A. D. Sarma, S. Bhandary, S. Bhattacharyya, A. Cangi & V. Aggarwal, A shallow hybrid classical–quantum spiking feedforward neural network for noise-robust image classification. Applied Soft Computing, 136, 110099, 2023. https://doi.org/10.1016/j.asoc.2023.110099.
[33] S. F. Stefenon, K. C. Yow, A. Nied & L. H. Meyer, Classification of distribution power grid structures using inception v3 deep neural network. Electrical Engineering, 104(6), 4557-4569, 2022. https://doi.org/10.1007/s00202-022-01641-1.
[34] S. F. Stefenon, G. Singh, K. C. Yow, & A. Cimatti, Semi-ProtoPNet deep neural network for the classification of defective power grid distribution structures. Sensors, 22 (13), 4859, 2022. https://doi.org/10.3390/s22134859.