Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

Improving Vegetation Encroachment Detection in Powerline Areas Using EfficientNet-Based U-Net Semantic Segmentation

Alissa Velia Royhatul Jannah[1], Nanik Suciati*[2]

[1,2]Department of Informatics, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

Email: [2]nanik@its.ac.id

Received : Apr 16, 2026; Revised : Apr 21, 2026; Accepted : Apr 26, 2026; Published : Apr 28, 2026

Abstract

Vegetation growing beyond safe limits has the potential to pose a threat to safety and the reliability of overhead powerlines, as well as cause financial losses for infrastructure providers. Identifying potential obstructions to overhead powerlines is crucial for addressing these issues. This study proposes the EFF-UNET semantic segmentation technique on the VEPL dataset to identify areas of overlap between vegetation and overhead powerlines by overlaying the two models. Visually, overhead powerlines have a thin pixel structure and are difficult to distinguish from the background or vegetation, whereas the feature extraction process in the U-Net encoder can degrade small objects due to progressive resolution loss. Modifications to the encoder in the baseline U-Net architecture utilize the EfficientNet family by comparing variants B0 through B7 to produce the best model. EfficientNet specifically employs compound scaling to optimize the network's resolution, depth, and width during feature extraction, thereby preserving information integrity during downsampling. Experimental results demonstrate the superiority of EfficientNetB7 through a measured trade-off compared to other models, where for vegetation segmentation, this model achieves an IoU of 0.9824, Accuracy of 0.9905, Dice of 0.9911, and Loss of 0.0089. Meanwhile, for powerline segmentation, the results show an IoU of 0.9153, Accuracy of 0.9978, Dice of 0.9558, and Loss of 0.0442. Based on these findings, EFF-UNET model successfully addresses the shortcomings of conventional models in preserving feature representation. This model is capable of improving the performance of vegetation and overhead powerlines segmentation to produce precise encroachment areas, thereby enabling accurate on-site infrastructure inspections.

Keywords : *EfficientNet, Encroachment Area Detection, Powerline Segmentation, Semantic Segmentation, Vegetation Monitoring.*

1. INTRODUCTION

Technological advancements in the digital age have transformed various aspects of human life and had an impact across numerous fields [1]. Human activities are now inseparable from the use of technology, such as the use of cell phones, laptops, and computers to access the internet [2]. The availability of electrical power and wireless telecommunications services has become an indispensable part of daily life, forming the foundation for nearly all economic, social, and security activities [3]. The reliability of overhead cable infrastructure is crucial for reaching end-users without disruption [4]. However, its distribution topology is vulnerable to environmental threats, particularly from vegetation that comes into contact with overhead cables [5]. The growth of tree canopies and branches along utility corridors poses a critical risk that can lead to failures in the electrical power distribution system [6],[7]. Direct contact between cables and vegetation, especially during extreme weather such as high winds, risks causing large-scale power outages, threatening public safety, and resulting in economic losses due to service interruptions [8],[9]. To mitigate these risks, utility companies proactively and routinely inspect their networks and maintain safe clearance distances between cables and vegetation [6].

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

Currently, traditional inspections still rely on ground patrols involving direct visual inspections in the field, thereby exposing technicians to significant geographical and safety risks [6], [10]. To address these challenges, technological innovations in Unmanned Aerial Vehicles (UAV) have revolutionized the monitoring of electrical infrastructure [11]. UAV equipped with various sensors, particularly RGB (Red, Green, Blue) cameras, provide a solution for monitoring overhead power line corridors and generating high-resolution imagery [6]. This RGB imagery also facilitates data analysis and updates [4], [6]. However, manually analyzing high-resolution imagery to identify powerline and vegetation encroachment remains a major challenge [12]. Several studies have demonstrated the effectiveness of deep learning algorithms for tree classification, safe distance monitoring, and cable detection in utility transmission corridors [13]–[16]. The model's adaptability makes artificial intelligence a solution for the initial assessment of electrical asset management [17].

The VEPL-NET method [18] was designed for semantic segmentation using a VGG16+U-Net architecture to detect encroachment areas between overhead powerline and vegetation using the VEPL dataset. The model was trained separately to produce vegetation segmentation and overhead power line segmentation. The segmentation results of both were overlaid to obtain the encroachment area, achieving an Intersection over Union (IoU) score of 0.64 for powerline detection and 0.7 for vegetation detection. Using the same dataset, the transformer-based SegFormer method was introduced in the study [17] to detect encroachment areas through an overlapping technique between the vegetation model and the powerline model, achieving an Intersection over Union (IoU) score of 0.86 for powerline detection and 0.94 for vegetation detection. In the VEPL dataset, powerlines visually have a thinner pixel structure compared to other objects. The challenge posed by U-Net repetitive feature extraction operations in the U-Net encoder can be detrimental when dealing with small objects, meaning that high-resolution details of overhead power lines are likely to be lost during the downsampling process [19],[20].

Several studies propose modifications to the U-Net to preserve both high-level semantic information and low-level spatial details. The Multi-Branch EfficientNet-U-Net, utilizing the EfficientNetB0 backbone, effectively extracts multiscale features in fine-scale structures and yields cleaner boundaries than the conventional U-Net, achieving a Dice coefficient of 0.93 [21]. A modified U-Net with an EfficientNet encoder detects ambiguous water boundaries in satellite imagery; hierarchical feature extraction from EfficientNetB1 to B5 substantially reduces interference from complex backgrounds [22]. EfficientNet encoders in hybrid architectures like Efficient R2U-Net and Dense U-Net for retinal blood vessel extraction address information loss in pooling layers [23]. Incorporating EfficientNet in U-Net not only improves spatial precision metrics such as Intersection over Union (IoU), but also enhances computational stability during network convergence [24]. fficientNet's compound scaling method balances resolution, width, and depth, enabling rich feature extraction without excessive computational cost [25]. This review confirms that EfficientNet encoders in U-Net effectively address the loss of thin structural details in minority classes across various cases.

Although several studies have successfully identified encroachment areas using a two-model overlapping technique, a limitation of previous research is that overhead powerlines visually exhibit a thin pixel structure and are difficult to distinguish from the background or vegetation, making evaluation results dependent on the number of pixels. This is evidenced by lower evaluation results for power line segmentation compared to vegetation segmentation. Previous research focused on the success of detecting encroachment areas, whereas the ability to preserve spatial details of features in powerline segmentation is crucial for precisely determining encroachment areas.

To address the issue of overhead powerline segmentation in the UAV VEPL dataset, an architecture is needed that can preserve spatial details to prevent the loss of overhead cable information due to downsampling. Inspired by several previous studies, structural modifications to the U-Net encoder are essential. This study proposes a modification of the EfficientNet architecture that offers a

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

compound scaling method to provide rich feature extraction without imposing excessive computational burden. This study aims to demonstrate that using EfficientNet as an encoder can enhance the ability to preserve fine details in minority objects. Several variants of EfficientNet were implemented to determine the optimal performance in preserving spatial details, particularly for overhead cables. Consequently, success in improving segmentation performance can lead to more precise detection of encroachment areas and maximize the monitoring of vegetation and overhead cable areas.

2. METHOD

This research method is designed to improve model performance in the task of semantic segmentation of overhead cables and vegetation using binary segmentation techniques, as shown in Figure 1. This study consists of several stages. First, to identify the most effective augmentation, we began by comparing two types of datasets that had been augmented geometrically and spectrally on the VEPL dataset. Next, the model is trained using a U-Net architecture with several variants of EfficientNet as the encoder to determine the model that best preserves spatial feature details. Model performance is evaluated using the Accuracy, IoU, Dice, and Loss metrics. Once the best model is identified, the two models are overlaid to identify areas of encroachment between vegetation and powerline.

2.1. DATASET

This study uses the VEPL dataset, which consists of high-resolution images collected along the Envigado road in Colombia [6]. This dataset is publicly available and focuses on vegetation segmentation in overhead power line corridors using drone imagery. Consisting of RGB pairs and masks for 532 images sized 256 x 256 pixels, this dataset provides labeled segment masks for vegetation (0, 255, 0), overhead powerline (110, 110, 110), and background (0, 0, 0). The availability of these multi-class segmentation masks is crucial for training and evaluating the proposed semantic segmentation model. Therefore, this dataset has the potential to serve as a standard benchmark for comparing the performance of the proposed method with state-of-the-art (SOTA) methods. The dataset has been augmented with geometric and spectral augmentations. Geometric augmentation is performed to alter the position and orientation of image objects so that the neural network becomes agnostic to such changes, using techniques such as random rotations, horizontal flips, and grid distortion with 3.724 total dataset. Meanwhile, spectral augmentation aims to enhance the model's robustness against variations in lighting and color by applying transformations such as random brightness and contrast, hue saturation, and CLAHE with 3.724 total dataset.

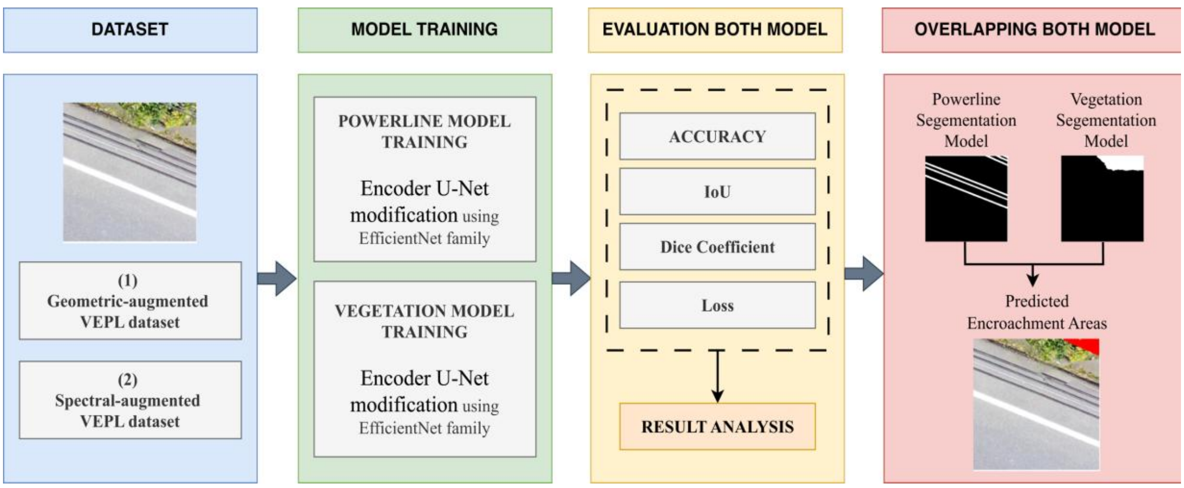

Figure 1. Methodology

Jurnal Teknik Informatika (JUTIF) Vol. 7, No. 2, April 2026, Page. 2023-2037
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

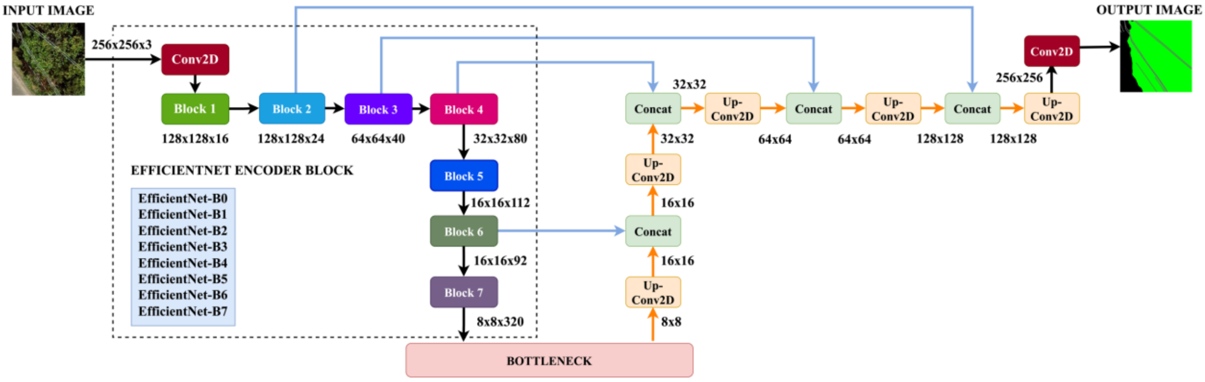

Figure 2. EFF-UNET architecture

2.2. MODEL ARCHITECTURE

To improve the performance of overhead power line and vegetation segmentation, this study proposes an EFF-UNET, as shown in Figure 2, to detect encroachment areas between vegetation and powerline. The research and evaluation process involves several sequential steps. First, the input images start with a resolution of 256 × 256 × 3 from the VEPL dataset. During the feature extraction stage, the input image resolution will be adjusted according to the variant of EfficientNet used, namely EfficientNetB0 through EfficientNetB7, as a replacement for the U-Net encoder. In this process, the input image resolution will be gradually reduced to evaluate the feature extraction process optimally. The downsampled images at each resolution are passed through blocks 2, 3, 4, and 6 using skip connections and then reconstructed via the decoder. After training is complete, the best weights obtained are used to generate predictions at high resolution, which are then evaluated.

2.3. EFF-UNET ENCODER MODIFICATION

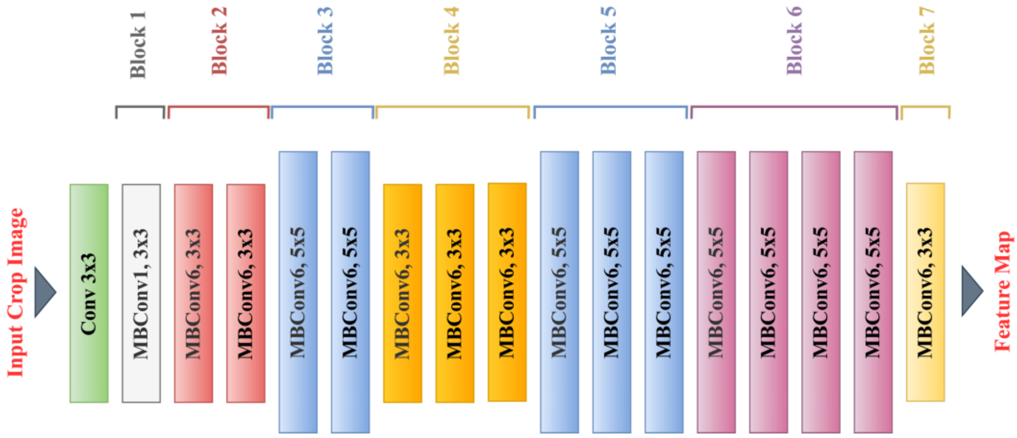

Figure 3. EfficientNet encoder block

In Figure 3, the U-Net encoder EfficientNet is responsible for preserving fine spatial details by processing the input image through a series of conv2D convolutional layers from block 1 to block 7, as shown in Eq. (1). Feature extraction is performed progressively by reducing the spatial dimensions of the features from 128x128 to the lowest resolution of 8x8. In this study, several skip connections in Eq. (2) are used to preserve the information from the feature extraction that has been downsampled.

$$j \in \{1, 2, 3, 4, 5, 6, 7\} \qquad (1)$$

$$S_k \in \{2, 3, 4, 6\} \qquad (2)$$

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

Table 1. Training Parameter

Feature	Parameter
Epoch	50
Batch Size	8
Learning Rate	0,0001
Optimizer	Adam
Image Size	[256,256]
Loss Function	Dice

The convolution process of the initial image (X) which has specific spatial dimensions and a certain number of channels. In the main processing stage, the seven blocks use the basic Mobile Inverted Bottleneck Convolution (MBConv) structure in iteration-i with kernel size variations $H \times W \times C$ as in Eq. (3). These are the fundamental blocks of EfficientNet that utilize Depthwise Separable Convolution and the Squeeze-and-Excitation (SE) mechanism to extract complex features from the image. This network is optimized using compound scaling based on resolution to scale the spatial dimensions $H \times W$, depth (d) to scale the number of layers, and channel width (w) to scale the number of channels or feature thickness to generate spatial features at the bottleneck (F_{enc}) which is the weighted sum of convolutions from all encoder layers as a weighted sum repeated as many times as d . L the number of layers, along with skip connections (S_k). Processing begins at Block 1 with a single MBConv1 (3x3) layer and proceeds to block 7, utilizing an MBConv6 layer with a combination of 3x3 and 5x5 kernel sizes. Processing begins at block 1 with a single MBConv1 layer (3x3) and continues through block 7, utilizing an MBConv6 layer with a combination of 3x3 and 5x5 kernel sizes.

$$F_{enc,}\{S_k\} = \sum_{j=1}^{d\,.\,L} MBConv_j \;\left(X_{(r\,.\,H)\times(r\,.\,W)\times(w\,.\,C)}\right) \quad (3)$$

Next, the set of spatial features extracted by the encoder and passed to the decoder via skip connections is defined as the set in Eq. (4) and is ready to be reconstructed to match the ground truth.

$$S = \{F_{enc\{S2\},}F_{enc\{S3\},}F_{enc\{S4\},}F_{enc\{S6\},}\} \quad (4)$$

2.4. TRAINING CONFIGURATION

The design of the EFF-UNET model for vegetation and powerline image segmentation was run in a Kaggle Notebook virtual environment on an NVIDIA Tesla P100 GPU with 16 GB of GDDR5X VRAM, supported by an Intel (R) Xeon (R) processor running at 2.20 GHz and 16 GB of RAM. The supporting software used was the Python programming language with libraries employed in this study, namely TensorFlow, Keras, Scikit-Learn, and Matplotlib. The network was trained using the Adam optimizer with a learning rate of 0.0001. For initialization, pre-trained weights from EFF-UNET were used for the unfrozen encoder section, with an input image size of 256 x 256 pixels. The training process runs for up to 50 epochs with an early stopping mechanism and uses dice as the loss function as shown in Table 1. The data split composition is 70% training data, 15% validation data, and 15% test data, where the final model is selected based on the best performance achieved during the validation phase.

2.5. EVALUATION METRIX

Each segmentation model was evaluated using a confusion matrix to compare the model’s predictions with the ground truth based on the classes of power lines, vegetation, and background, using the corresponding true positive (TP), true negative (TN), false positive (FP), and false negative (FN)

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

values. Model testing was conducted by dividing the random sampling data with a random state (42) into 2,606 training images (70%), 559 validation images (15%), and 559 test images (15%).

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (5)$$

$$IoU = \frac{TP}{TP+FP+FN} \quad (6)$$

$$Dice\ Coefficient = \frac{2\ .\ TP}{2\ .\ TP+FP+FN} \quad (7)$$

$$Loss = 1 - Dice\ Coefficient \quad (8)$$

Accuracy is the overall percentage of correct predictions, which measures how often the model successfully classifies pixels as positive or negative using a formula such as that in Eq. (5). IoU is a metric for segmentation because it explicitly accounts for spatial accuracy and object localization using a formula such as that in Eq. (6). Dice Coefficient is an evaluation metric widely used in semantic segmentation to measure the similarity between the model's prediction results and the ground truth, using a formula such as that in Eq. (7). Finally, Loss is the primary loss function optimized during the training process, using a formula such as that in Eq. (8).

2.6. OVERLAP BOTH MODEL

The overlapping technique in this system operates based on the principle of logical intersection to determine encroachment areas as shown in Figure 4. First, the system uses two state-of-the-art segmentation models, EFF-UNET, which have been trained separately on the same images to determine the positions of powerline and vegetation. After both models generate their respective detection maps, the system stacks and aligns the two maps. An area is only designated as an encroachment area if it is detected as a power line and as vegetation at the same pixel. In the final visualization, these overlapping points are then marked in red on the original image.

$$Encroachment_{(x,y)} = Vegetation_{(x,y)} \times Powerline_{(x,y)} \quad (9)$$

This slicing operation is performed on a pixel-by-pixel basis. Binary maps from EFF-UNET predictions assign a value of 1 to represent $Vegetation_{(x,y)}$ or $Powerline_{(x,y)}$ objects and 0 for the background. Next, pixel values at coordinates (x, y) are processed as shown in Eq. (9). The Encroachment binary map $Encroachment_{(x,y)}$ is set to 1 only if both vegetation and powerline maps have a value of 1 at the same coordinates, ensuring detection occurs only at precise spatial overlaps.

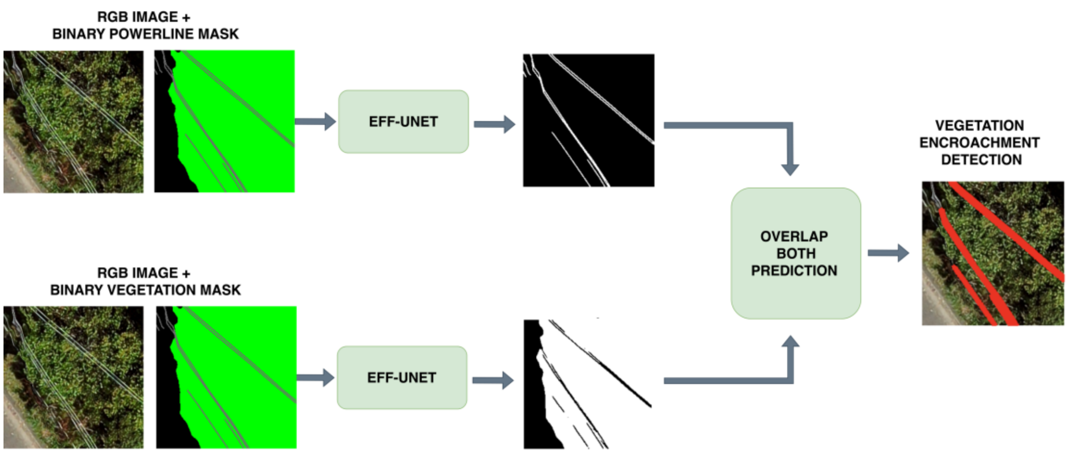

Figure 4. Overlap both model

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

3. RESULT

3.1. ABLATION STUDY

Ablation studies were conducted to investigate and validate the individual contributions of various proposed modules within the model architecture. This evaluation was performed by systematically replacing the encoder architecture in U-Net to observe its impact on the model's overall performance metrics. To analyze the influence of each model component on segmentation performance, we conducted an ablation study with several variants of the U-Net encoder. Experimental consistency was maintained by testing on the VEPL dataset, which had been augmented with spectral and geometric data. We used the Dice coefficient as the loss function. The dataset was split 70:15:15 for training, validation, and testing. Model training was performed over 50 epochs using the Adam optimizer and a learning rate of 0.0001. This analysis is crucial for providing empirical evidence. Thus, we demonstrate improved segmentation metrics on thin structured minority classes, particularly overhead cables, which are directly attributable to the feature retention capabilities of the proposed encoder, thereby justifying the complexity of the chosen architectural design.

3.2. TEST SET PERFORMANCE EVALUATION ENCODER EFFICIENTNET FAMILY

To analyze the impact of each model component on segmentation performance, we tested several variants of the U-Net encoder with the following backbones: EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7.

Table 2. Vegetation model evaluation

Encoder	Augmen-tation	Vegetation				Powerline			
		IoU	Acc	Dice	Loss	IoU	Acc	Dice	Loss
Efficient-NetB0	Geometric	0.9193	0.9577	0.9580	0.0420	0.4775	0.9810	0.6464	0.3536
	Spectral	0.9759	0.9879	0.9879	0.0122	0.8664	0.9964	0.9284	0.0716
Efficient-NetB1	Geometric	0.9279	0.9627	0.9626	0.0374	0.4889	0.9810	0.6568	0.3432
	Spectral	0.9765	0.9882	0.9881	0.0119	0.8791	0.9965	0.9356	0.0644
Efficient-NetB2	Geometric	0.9326	0.9646	0.9651	0.0349	0.5084	0.9815	0.6740	0.9815
	Spectral	0.9782	0.9889	0.9890	0.0110	0.8806	0.9969	0.9365	0.0635
Efficient-NetB3	Geometric	0.9379	0.9678	0.9679	0.0321	0.5103	0.9813	0.6758	0.3242
	Spectral	0.9818	0.9901	0.9908	0.0092	0.8821	0.9967	0.9373	0.0627
Efficient-NetB4	Geometric	0.9257	0.9622	0.9614	0.0386	0.4986	0.9815	0.6655	0.3345
	Spectral	0.9798	0.9892	0.9898	0.0102	0.8963	0.9972	0.9453	0.0547
Efficient-NetB5	Geometric	0.9345	0.9659	0.9661	0.0339	0.5325	0.9830	0.6949	0.3051
	Spectral	0.9787	0.9895	0.9892	0.0108	0.9053	0.9973	0.9503	0.0497
Efficient-NetB6	Geometric	0.9387	0.9687	0.9684	0.0316	0.5229	0.9826	0.6867	0.3133
	Spectral	0.9813	0.9901	0.9905	0.0095	0.9052	0.9975	0.9502	0.0498
Efficient-NetB7	Geometric	0.9375	0.9677	0.9677	0.0323	0.5407	0.9833	0.7019	0.2981
	Spectral	**0.9824**	**0.9905**	**0.9911**	**0.0089**	**0.9153**	**0.9978**	**0.9558**	**0.0442**

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

To analyze the impact of each model component on segmentation performance, we tested several variants of the U-Net encoder with the following backbones: EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7. This evaluation was designed to objectively measure the model's generalization ability when faced with UAV imagery using test data that had never been seen during either the training or validation phases. The performance of the modified U-Net architecture with EfficientNet encoders was assessed based on strict pixel segmentation metrics, including Intersection over Union (IoU), F1-Score, Precision, and Recall. The analysis in this section aims to demonstrate the model's reliability in handling extreme class imbalance in aerial cable objects, while also proving its robustness against visual background variations that represent real-world conditions.

3.2.1. VEGETATION SEGMENTATION MODEL EVALUATION

For the vegetation segmentation task, the model was trained using several training schemes involving two types of data augmentation and various EfficientNet variants as encoders. The goal was to identify the optimal model performance. Evaluation of the vegetation class confirmed the performance of the proposed hybrid architecture, as shown in Table 2. The EFF-UNET architecture with an EfficientNetB7 encoder combined with spectral augmentation outperformed other models, achieving an IoU of 0.9824, an accuracy of 0.9905, a Dice score of 0.9911, and a loss of 0.0089. The high accuracy confirms the model's reliability in separating vegetation pixels from background elements. The high IoU metric confirms the similarity between the predicted vegetation areas and the ground truth. These four metrics comprehensively represent the quality of the vegetation visual segmentation output well. At the architectural level, the network depth in EfficientNetB7 enables hierarchical feature extraction, allowing it not only to recognize vegetation pixel clusters globally but also to precisely delineate complex edge boundaries. In the context of infrastructure monitoring, the reliability of mapping vegetation outer boundaries serves as the primary foundation for identifying encroachment areas when overlaps occur.

3.2.2. POWERLINE SEGMENTATION MODEL EVALUATION

The evaluation of the overhead powerline class shown in Table 2 indicates that the EFF-UNET architecture, featuring an EfficientNetB7 encoder combined with spectral augmentation, outperformed other models with an IoU of 0.9153, an accuracy of 0.9978, a Dice score of 0.9558, and a loss of 0.0442. The high IoU is used to measure the similarity between the predicted powerline area and the ground truth. This confirms that the model's predictions are precise and do not suffer from dimensional dilation, ensuring that cable thickness estimates remain proportional. The optimization of compound scaling in the resolution and depth of the EfficientNetB7 network has proven capable of preserving the integrity of fine details and maintaining the continuity of cable pixels during the downsampling process. This

Table 3. Parameter comparison

Model	Input Size	Total params	Trainable params	Non-trainable params
EfficientNetB0	224x224	40.807.364	40.760.669	46.695
EfficientNetB1	240x240	43.333.032	43.266.305	66.727
EfficientNetB2	260x260	46.038.170	45.965.731	72.439
EfficientNetB3	300x300	50.758.864	50.666.313	92.551
EfficientNetB4	380x380	60.783.360	60.652.521	130.839
EfficientNetB5	456x456	223.416.636	74.412.625	178.759
EfficientNetB6	528x528	270.138.276	89.969.113	230.935
EfficientNetB7	600x600	348.946.460	116.209.585	317.703

Jurnal Teknik Informatika (JUTIF) Vol. 7, No. 2, April 2026, Page. 2023-2037
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

success has significant implications for infrastructure monitoring, the reliability of powerline mapping also serves as a key foundation for identifying encroachment areas when overlaps occur.

3.2.3. PARAMETER COMPARISON BETWEEN EFFICIENTNET FAMILY (B1-B7)

The combined U-Net and EfficientNet variants require the computational resources shown in Table 3 to determine the best model. The architectural progression from EfficientNetB0 to EfficientNetB7 was designed using the principle of compound scaling, which simultaneously increases the input image resolution, network depth, and channel width. Feature extraction is performed by systematically increasing the input size and the number of parameters to achieve more complex results. EfficientNetB0 is designed as the lightest model with a standard input resolution of 224 × 224 pixels and the lowest number of trainable parameters. As the variants scale up to EfficientNetB7, the input resolution increases to 600 × 600 pixels, with an exponential rise in the number of network parameters. The correlation among the EfficientNet variants indicates that the lighter the model, the more efficient the inference time and computational load. EfficientNet variants with low computational load are ideal for real-time inference processing on UAV devices. Conversely, higher computational capacity, such as in EfficientNetB7, has proven to offer the advantage of a wide receptive field for extracting spatial details, thereby preserving fine details suitable for analyzing complex and small objects.

3.2.4. EFF-UNET

The following graph shows the training performance of the best model using EfficientNetB7 on a U-Net encoder at 50 epochs. IoU measures how precisely the model's predicted areas for powerline segmentation overlap with the ground truth. In Figure 5, the graph shows a stable curve around 0.9, although there was a slight decline between epochs 20 and 30. The accuracy curve spikes very high at the beginning of the epochs and settles at a value close to 0.99. After epoch 30, the validation loss slowly drops again and stabilizes close to the training loss. The red line consistently overlaps with the yellow line, indicating that the model is learning well.

The IoU and Dice Coefficient plots are the most critical indicators for evaluating how well the predicted vegetation areas match the actual ones. Figure 6, show that the model recovered quickly, gradually climbed back up, and eventually stabilized near the training line with a score above 0.9. This indicates that the vegetation mask predictions are highly accurate.

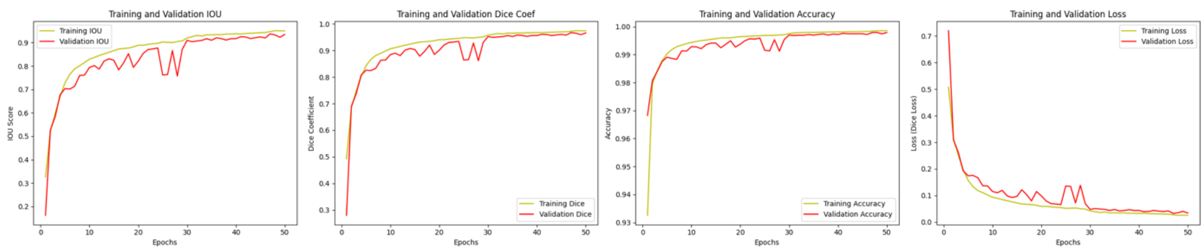

Figure 5. Training validation powerline

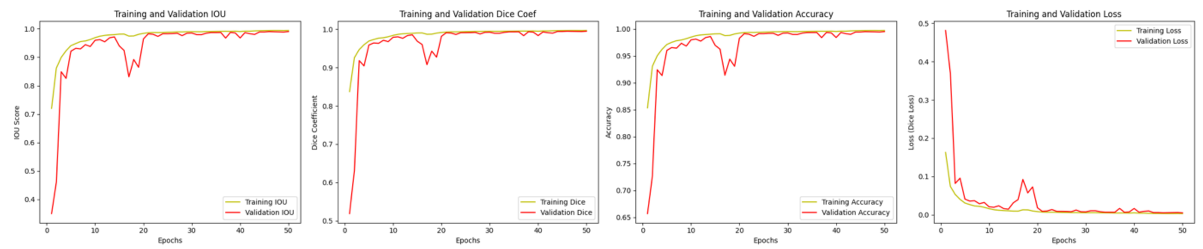

Figure 6. Training validation vegetation

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

The accuracy graph shows a trend consistent with values nearly reaching 1.0. Fluctuations in the middle of the epoch are also visible here. In pixel segmentation, high accuracy should be approached with caution if the IoU is low, as the model may simply be good at guessing the background color. In vegetation segmentation, IoU and Dice are proven to be high; accuracy is nearly perfect due to the model's ability to accurately detect vegetation classes. The loss curve provides the strongest validation that the model is not overfitting. The training loss curve consistently decreases until it approaches zero. The combination of EfficientNetB7 and U-Net results in a very large, deep, and complex feature extractor model. The training plots demonstrate that the model learns effectively, capturing complex visual features of vegetation and successfully applying them to new validation data without losing accuracy.

4. DISCUSSIONS

This study has successfully applied and evaluated deep learning techniques for semantic image segmentation in the maintenance of overhead cable infrastructure in vegetated areas. A modified U-Net encoder, utilizing various variants of EfficientNet, was compared to achieve the best results when segmenting two types of areas: overhead cable routes and vegetated areas.

4.1. EFFICIENTNET ENCODER EFFECTIVENESS

Modifications to U-Net using the EfficientNet encoder were compared across variants from EfficientNetB0 to EfficientNetB7, with training environments standardized to ensure a fair final evaluation. Test results demonstrate that the EFF-UNET model with EfficientNetB7 outperforms other EfficientNet variants. Progressive model evaluation on a spectral vegetation augmentation dataset showed an improvement from EfficientNetB0 to EfficientNetB3, reaching an IoU of 0.9818. However, performance dropped when using the EfficientNetB4 model with an IoU of 0.9798 and the EfficientNetB5 model with an IoU of 0.9787. Finally, EfficientNetB7 demonstrated an improvement across all evaluation metrics. This demonstrates that the use of EfficientNetB7, with its high computational capacity, effectively expands the receptive field to extract spatial details. The EFF-UNET model successfully preserves features of complex objects and is capable of delineating area boundaries. Consequently, this model enhances the performance of vegetation segmentation and powerline segmentation to produce precise encroachment areas when the two models are overlaid.

4.2. AUGMENTATION ANALYSIS

The evaluation results show that the performance of models using spectral augmentation outperforms that of models using geometric augmentation in every EfficientNet test. Spectral augmentation, which applies transformations such as random brightness and contrast, hue and saturation, and CLAHE, successfully improves the model's robustness to variations in lighting and color. This approach preserves image features intact while simulating real-world visual challenges, such as tree shadow occlusion and light exposure fluctuations in UAV imagery. Since spatial integrity remains intact, the Squeeze-and-Excitation (SE) mechanism in EfficientNet can operate optimally to focus on persistent powerline textures and complex vegetation textures. This enables the network to learn to extract features that are robust against lighting noise without sacrificing thin structural details. This combination not only prevents feature erosion in deep convolutional layers but also produces an EFF-UNET segmentation model that is robust against visual degradation in real-world environments. This provides a precise algorithmic foundation for detecting powerline infrastructure encroachment areas.

Vol. 7, No. 2, April 2026, Page. 2023-2037
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

Table 4. Model comparison

Encoder	Augmentation	IoU	Accuracy	Dice	Loss
		Vegetation			
U-Net [18]	Geometric	0.9038	0.9508	0.9495	0.0505
	Spectral	0.8894	0.9429	0.9415	0.0585
VGG16+U-Net [18]	Geometric	0.5069	0.5151	0.6728	0.3272
	Spectral	0.965	0.9807	0.9822	0.0178
Meta-Learner [17]	Geometric	0.8136	0.8968	0.8972	0.1028
	Spectral	0.5152	0.5239	0.6801	0.3199
EfficientNetB0	Geometric	0.9193	0.9577	0.958	0.042
	Spectral	0.9759	0.9879	0.9879	0.0122
EfficientNetB7	Geometric	0.9375	0.9677	0.9677	0.0323
	Spectral	**0.9824**	**0.9905**	**0.9911**	**0.0089**
		Powerline			
U-Net [18]	Geometric	0.5406	0.983	0.7018	0.2982
	Spectral	0.7957	0.9939	0.8862	0.1138
VGG16+U-Net [18]	Geometric	0.561	0.9844	0.7187	0.2813
	Spectral	0.8405	0.9951	0.9133	0.0867
Meta-Learner [17]	Geometric	0.3234	0.9723	0.4887	0.5113
	Spectral	0.5087	0.9827	0.6743	0.3257
EfficientNetB0	Geometric	0.4775	0.981	0.6464	0.3536
	Spectral	0.8664	0.9964	0.9284	0.0716
EfficientNetB7	Geometric	0.5407	0.9833	0.7019	0.2981
	Spectral	**0.9153**	**0.9978**	**0.9558**	**0.0442**

4.3. COMPARISON WITH SOTA

Table 4 shows a comparison between EfficientNetB0, the lightest EfficientNet model, and EfficientNetB7, the model with the highest computational cost. This comparison employs state-of-the-art (SOTA) methods by comparing EFF-UNET with previous research on vegetation and overhead power line segmentation. The reference models are the baseline U-Net, VGG16+U-Net, and MetaLearner. These methods were selected to validate the reliability and effectiveness of this approach, as well as to objectively map the contribution of EFF-UNET in the literature. The evaluation results show that the EfficientNetB7 model successfully improved evaluation metrics using spectral segmentation, with an IoU of 0.9824, an accuracy of 0.9905, Dice 0.9911, and Loss 0.0089 for vegetation segmentation, as well as an IoU of 0.9153, an accuracy of 0.9978, Dice 0.9558, and Loss 0.0442 for overhead cable segmentation.

4.4. IMPROVEMENT ANALYSIS

The model's performance was also compared to architectures from previous studies, namely the baseline U-Net and VGG16+U-Net, through retraining and configuration alignment. Overall, the experimental results show that EFF-UNET, with the EfficientNetB0 backbone as the lightest model, outperforms the baseline U-Net (+7.98%) and VGG16+U-Net (+1.13%) in vegetation segmentation. In powerline segmentation, EFF-UNET also outperformed the baseline U-Net (+8.89%) and VGG16+U-Net (+3.08%). Meanwhile, computationally intensive models such as EfficientNetB7 outperform the U-Net baseline (+8.70%) and VGG16+U-Net (+1.80%) in vegetation segmentation, as well as the U-Net baseline (+15.03%) and VGG16+U-Net (+8.90%) in powerline segmentation. The performance of the standard U-Net architecture indicates that the use of a simple encoder relying solely on a stack of

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

standard convolutional layers is not yet capable of optimally extracting complex spatial features on the dataset used. The performance improvement achieved using VGG16+U-Net outperforms the standard U-Net, enabling the model to capture a richer feature hierarchy. However, the efficiency and generalization capabilities of VGG16+U-Net still fade behind those of compound scaling-based architectures. The architectural approach using EfficientNet as the encoder consistently delivers the most substantial improvements across various target classes. This advantage is driven by the Squeeze-and-Excitation (SE) mechanism and parameter optimization via compound scaling in EfficientNet, which enables the model to adaptively rebalance feature channels. These results demonstrate that using the EfficientNet architecture as a U-Net encoder successfully extracts features and significantly enhances U-Net's ability to perform more precise and robust image segmentation compared to conventional approaches.

4.5. MODEL VISUALIZATION

Figure 7 presents representative segmentation results by comparing predictions from the three architectures using spectral augmentation techniques. Each model generates two predictions, one for vegetation and one for powerline, which are then overlapped to detect encroachment areas. Each prediction is compared with the ground truth mask. The visualization shows that all models successfully detected encroachment areas, with EfficientNetB7 demonstrating better preservation of detail in encroachment areas compared to the other models.

5. CONCLUSION

This study successfully evaluated deep learning techniques using EFF-UNET for semantic segmentation tasks in detecting encroachment areas. Comparative analysis reveals a trade-off between model performance and computational complexity. EfficientNet, with its lightest model such as EfficientNetB0, has outperformed the U-Net and VGG16+U-Net baselines and offers an ideal solution for computationally intensive tasks, such as real-time inference processing on UAVs.

Meanwhile, EfficientNetB7, as the model with the highest computational cost among the EfficientNet variants, outperforms the U-Net and VGG16+U-Net baselines, as well as other EfficientNet variants, in addressing the challenge of preserving the spatial details of small objects. The EfficientNetB7 architecture, serving as a U-Net encoder, successfully improves evaluation results, particularly for the powerline class, and addresses class imbalance caused by thin and linear topologies.

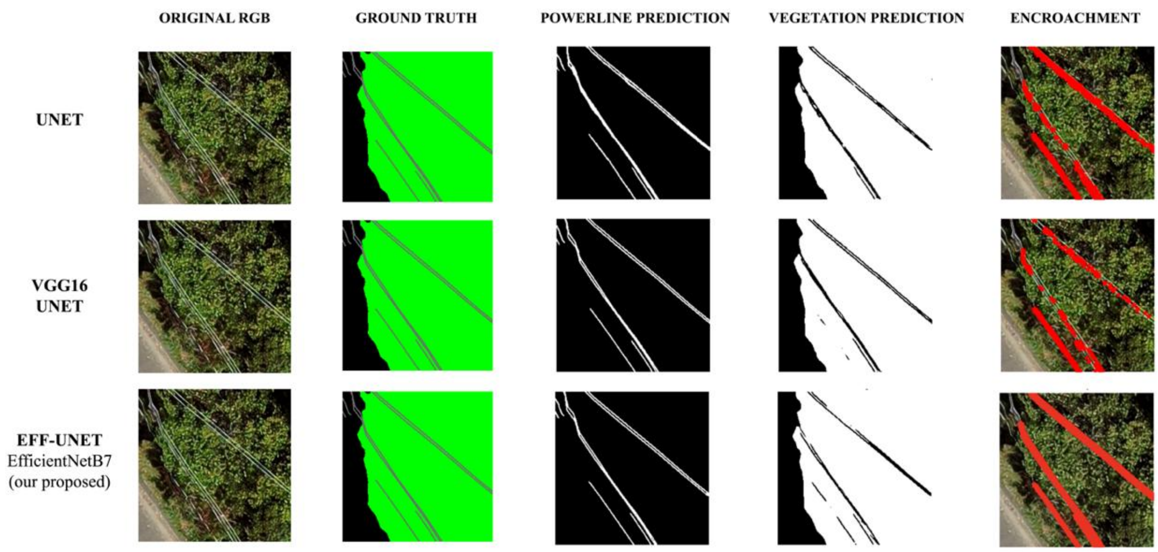

Figure 7. Model comparison

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5863

EfficientNetB7, with its greater computational capacity, offers the advantage of a wider receptive field for extracting spatial details, thereby successfully preserving features suitable for analyzing complex objects and delineating area boundaries.

The model's performance with spectral augmentation successfully enhances the model's robustness against lighting and color variations, enabling the network to learn to extract features resistant to lighting noise without sacrificing thin structural details. This results in an EFF-UNET segmentation model that is robust against visual degradation in real-world environments and produces a precise segmentation model for accurately detecting powerline infrastructure encroachment areas.

Although the proposed model has successfully addressed the challenge of improving feature representation during the downsampling process, this study has a major limitation in the implementation of encroachment detection. First, the training cycle is doubled to extract vegetation and overhead cable classes. The use of the EfficientNetB7 model demands massive memory allocation and computational load due to the high number of trainable parameters. This dual-training paradigm not only doubles the computational and memory resource requirements but also burdens the architecture and requires longer inference times, which can limit the system's scalability. Furthermore, encroachment detection is not learned directly by the model using ground truth, making the model susceptible to optical illusions and perspective distortions. For example, a tree in the distant background may appear to overlap or touch a cable in the foreground. This has the potential to result in a high false positive rate in real-world encroachment detection. To address these operational and analytical limitations, future research needs to implement a Unified Multi-Task Learning architecture. This way, the model can efficiently segment vegetation and overhead cables and predict encroachment areas in a single propagation cycle, thereby significantly reducing computational load.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

ACKNOWLEDGEMENT

This research was supported by the Department of Informatics, Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia.

REFERENCES

[1] C. A. Dekker, S. E. Baumgartner, S. R. Sumter, and J. Ohme, "Beyond the Buzz: Investigating the Effects of a Notification-Disabling Intervention on Smartphone Behavior and Digital Well-Being," *Media Psychology*, vol. 28, no. 1, pp. 162–188, 2025, doi: 10.1080/15213269.2024.2334025.

[2] L. Kiburu, N. Njiraini, and N. Boso, "Social networks and consumer technology usage: A systematic literature review and future research directions," *Cogent Business and Management*, vol. 10, no. 1, 2023, doi: 10.1080/23311975.2022.2153487.

[3] M. A. Mahfud, D. M. A. Siregar, and M. F. Malik, "Nussbaum's Justice Theory Lens: the Digital Divide and Unfulfillment of the Right to Public Service in Traditional Society," *Media Iuris*, vol. 9, no. 1, pp. 97–128, 2026, doi: 10.20473/mi.v9i1.77214.

[4] F. Bahreini and A. Hammad, "Detection of Vegetation Proximity to Power Lines: Critical Review and Research Roadmap," *Forests*, vol. 16, no. 11, pp. 1–40, 2025, doi: 10.3390/f16111658.

[5] M. F. Pacevicius, M. Ramos, C. T. Eriksen, and N. Paltrinieri, "Data-Informed Risk Analysis of Power Grids: Application of Method for Managing Heterogeneous Datasets," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 11, no. 2, 2025, doi: 10.1115/1.4066257.

[6] M. Cano-Solis, J. R. Ballesteros, and J. W. Branch-Bedoya, "VEPL Dataset: A Vegetation Encroachment in Power Line Corridors Dataset for Semantic Segmentation of Drone Aerial Orthomosaics," *Data*, vol. 8, no. 8, 2023, doi: 10.3390/data8080128.

[7] V. S. Rajkumar, A. Stefanov, A. Presekal, P. Palensky, and J. L. R. Torres, "Cyber Attacks on Power Grids: Causes and Propagation of Cascading Failures," *IEEE Access*, vol. 11, no. July, pp. 103154–103176, 2023, doi: 10.1109/ACCESS.2023.3317695.

[8] S. Paul, A. Poudyal, S. Poudel, A. Dubey, and Z. Wang, "Resilience assessment and planning in power distribution systems: Past and future considerations," *Renewable and Sustainable Energy Reviews*, vol. 189, no. PB, p. 113991, 2024, doi: 10.1016/j.rser.2023.113991.

[9] R. Suttle, B. Kane, and D. Bloniarz, "Comparing the Structure, Function, Value, and Risk of Managed and Unmanaged Trees along Rights-of-Way and Streets in Massachusetts," *Forests*, vol. 13, no. 10, 2022, doi: 10.3390/f13101602.

[10] M. A. A. Faisal, I. Mecheter, Y. Qiblawey, J. H. Fernandez, M. E. H. Chowdhury, and S. Kiranyaz, "Deep learning in automated power line inspection: A review," *Applied Energy*, vol. 385, no. May 2024, p. 125507, 2025, doi: 10.1016/j.apenergy.2025.125507.

[11] Z. Ni *et al.*, "Research on UAV-LiDAR-Based Detection and Prediction of Tree Risks on Transmission Lines," *Forests*, vol. 16, no. 4, 2025, doi: 10.3390/f16040578.

[12] A. Kyuroson, A. Koval, and G. Nikolakopoulos, "Autonomous Point Cloud Segmentation for Power Lines Inspection in Smart Grid," *International Federation of Automatic Control*, vol. 56, no. 2, pp. 11754–11761, 2023, doi: 10.1016/j.ifacol.2023.10.562.

[13] B. Harini, G. Mp, and M. Manju, "AI-Driven Vegetation Monitoring and Fire Risk Mitigation for Wildfire Prevention and Powerline Safety," *Proceedings of 2025 3rd International Conference on Intelligent Systems, Advanced Computing, and Communication, ISACC 2025*, pp. 145–150, 2025, doi: 10.1109/ISACC65211.2025.10969250.

[14] A. Al-Najjar, M. Amini, S. Rajan, and J. R. Green, "Identifying Areas of High-Risk Vegetation Encroachment on Electrical Powerlines Using Mobile and Airborne Laser Scanned Point Clouds," *IEEE Sensors Journal*, vol. 24, no. 14, pp. 22129–22143, 2024, doi: 10.1109/JSEN.2023.3348785.

[15] Y. Zhou, Z. Feng, C. Chen, and F. Yu, "Bilinear Distance Feature Network for Semantic Segmentation in PowerLine Corridor Point Clouds," *Sensors*, vol. 24, no. 15, 2024, doi: 10.3390/s24155021.

[16] J. Li, H. Zheng, P. Liu, Y. Liang, F. Shuang, and J. Huang, "Safety monitoring method for powerline corridors based on single-stage detector and visual matching," *High Voltage*, vol. 9, no. 4, pp. 805–815, 2024, doi: 10.1049/hve2.12464.

[17] D. Ayobo-Abongo, W. Jaafar, and R. Langar, "A Novel Hybrid ML Approach for Powerline - Vegetation Encroachment Area Identification," *IEEE International Symposium on Industrial Electronics*, no. Ml, pp. 1–7, 2025, doi: 10.1109/ISIE62713.2025.11124746.

[18] M. Cano-Solis, J. R. Ballesteros, and G. Sanchez-Torres, "VEPL-Net: A Deep Learning Ensemble for Automatic Segmentation of Vegetation Encroachment in Power Line Corridors Using UAV Imagery," *ISPRS International Journal of Geo-Information*, vol. 12, no. 11, 2023, doi: 10.3390/ijgi12110454.

[19] C. Torres de Almeida *et al.*, "Canopy Height Mapping by Sentinel 1 and 2 Satellite Images, Airborne LiDAR Data, and Machine Learning," *Remote Sensing*, vol. 14, no. 16, pp. 1–21, 2022, doi: 10.3390/rs14164112.

[20] M. S. A. Mohd Rapheal *et al.*, "Machine learning approach for tenaga nasional berhad (tnb) overhead powerline and electricity pole inventory using mobile laser scanning data," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 46, no. 4/W3-2021, pp. 239–246, 2022, doi: 10.5194/isprs-archives-XLVI-4-W3-2021-239-2022.

[21] N. P. Husain, "A Multi-Branch EfficientNet-U-Net Hybrid Framework for Segmentation of Oyster Mushrooms in Cultivation Media," *Journal of Embedded System Security and Intelligent Systems*, vol. 7, no. 1, pp. 78–91, 2026, doi: 10.59562/jessi.v7i1.2607.

[22] M. Ezz, A. S. Alaerjan, A. M. Mostafa, N. Laban, and H. H. Zeyada, "Progressive Attention-Enhanced EfficientNet–UNet for Robust Water-Body Mapping from Satellite Imagery,"

Sensors, vol. 26, no. 3, pp. 1–25, 2026, doi: 10.3390/s26030963.

[23] N. Siddique, S. Paheding, A. A. Reyes Angulo, M. Z. Alom, and V. K. Devabhaktuni, “Fractal, recurrent, and dense U-Net architectures with EfficientNet encoder for medical image segmentation,” *Journal of Medical Imaging*, vol. 9, no. 06, pp. 1–22, 2022, doi: 10.1117/1.jmi.9.6.064004.

[24] S. Y. Lin and C. L. Lin, “Brain tumor segmentation using U-Net in conjunction with EfficientNet,” *PeerJ Computer Science*, vol. 10, 2024, doi: 10.7717/peerj-cs.1754.

[25] M. J. Mahoney, L. K. Johnson, A. Z. Guinan, and C. M. Beier, “Classification and mapping of low-statured shrubland cover types in post-agricultural landscapes of the US Northeast,” *International Journal of Remote Sensing*, vol. 43, no. 19–24, pp. 7117–7138, 2022, doi: 10.1080/01431161.2022.2155086.