# Automated Power Lines Vegetation Monitoring using High-Resolution Satellite Imagery

Michele Gazzea (iD), *Member, IEEE*, Michael Pacevicius, Dyre Oliver Dammann,
Alla Sapronova, Torleif Markussen Lunde, and Reza Arghandeh (iD), *Senior Member, IEEE*

*Abstract*—Vegetation Management is a significant preventive maintenance expense in many power transmission and distribution companies. Traditional Vegetation Management operational practices have proven ineffective and are rapidly becoming obsolete due to the lack of frequent inspection of vegetation and environmental states. The rise of satellite imagery data and machine learning provides an opportunity to close the loop with continuous data-driven vegetation monitoring. This paper proposes an automated framework for monitoring vegetation along power lines using high-resolution satellite imagery and a semi-supervised machine learning algorithm. The proposed satellite-based vegetation monitoring framework aims to reduce the cost and time of power line monitoring by partially replacing ground patrols and helicopter or drone inspection with satellite data analytics. It is implemented and demonstrated for a power distribution system operator (DSO) in the west of Norway. For further assessment, the satellite-based algorithm outcomes are compared with LiDAR survey data collected by helicopters. The results show the potential of the solution for reducing the monitoring costs for electric utilities.

*Index Terms*—Satellite imagery, Vegetation management, Power systems, Electric grid monitoring, Semi-supervised segmentation

## I. INTRODUCTION

**P**OWER transmission and distribution networks spread across countries and pass through forests, over various terrain, and cities on their journey to electricity consumers. Whenever vegetation interferes with power lines, it brings safety, economic, and environmental risks. Vegetation, combined with severe weather conditions, is the predominant reason for outages in power systems that put millions of people in darkness and bring billions of dollars in economic damage [1]. In areas with severe drought, vegetation encroachment in power lines' right-of-way (ROW) can cause massive wildfires with high fatality rate [2], [3]. Vegetation monitoring and management is becoming ever more important in the wake of climate change and the increasing frequency and duration of extreme weather events [4].

Utilities traditionally take a time-based approach with a fixed cycle to vegetation monitoring by sending ground-based patrol for visual line inspection and flying helicopters or drones for optical and LiDAR surveys of the power lines. Due to the vast size of service territories, the length of power lines, and the line inspection costs, vegetation monitoring's typical cycle varies between one to ten years for different electric utilities [5]. For example, the US electrical grid has more than 200,000 miles of high-voltage transmission lines and 5.5 million miles of local distribution lines [6].

Existing literature on vegetation monitoring mostly uses LiDAR surveys performed by helicopters or drones [7], [8]. LiDAR data provide an accurate 3D representation of an environment. However, LiDAR data acquisition and processing are extremely pricey and time-consuming. If LiDAR-based line monitoring is performed for a large transmission or distribution company, it is often done infrequently at an interval of once every 5 to 10 years to scan the whole service area [9].

In recent years, the drop in launching costs and the growing number of satellites and mini-satellites in orbit with high-quality sensors has significantly reduced the cost of satellite imagery [10]. Commercial satellite providers can offer high-resolution images (0.25 or 0.5 meters/pixel) with frequent revisiting time that covers most of the world. Consequently, it brings the opportunity to combine scale, frequency, and cost efficiency to enhance situational awareness regarding vegetation encroachment in power lines' right-of-way using high-resolution satellite imagery [11], [12]. Therefore, vegetation management can be changed from traditional time-based monitoring to risk-based monitoring.

Some studies [13], [14], [15] made use of multispectral stereo pairs of satellite images for each specific area to identify trees along power lines. However, stereo images are challenging to capture and are costly for large scale areas [16]. This paper proposes a machine learning-based algorithm for vegetation detection using a single satellite image, which is more cost-effective.

Vegetation detection from single monocular images is also a well-studied topic, particularly in the forest management of agriculture and urban areas using classic image processing tools for vegetation detection [17], [18], [19], [20]. However, such approaches have been developed to work well where trees are easily distinguishable, i.e., in low-density vegetation areas [21] or when the trees are regularly spaced from each other in orchids, which is not the case in the vast majority of power

lines' ROW [22].

Nowadays, Convolution Neural Networks (CNNs), have become the leading machine learning methodology in many fields due to their effectiveness at extracting feature representations from images for classification and segmentation purposes [23], [24]. For example, [25] proposed a semantic segmentation-based deep learning method to classify vegetation (tree, shrub, and grass) using only RGB images. In a similar work, [26] used a U-Net architecture for analyzing high-resolution satellite images to map forests. However, deep learning methods are generally supervised approaches and need massive labeled datasets for the training, which is extremely scant and expensive for satellite imagery and remote sensing applications. Weakly-supervised methods are, in general, more practical [27].

This paper proposes a framework to monitor vegetation proximity to power lines using high-resolution satellite images. From a methodological point of view, it is a semi-supervised approach for vegetation detection that is a combination of a deep unsupervised architecture and a supervised machine learning algorithm. Being unsupervised, the first layer of the proposed framework does not need any training data and takes advantage of deep learning to capture meaningful patterns in satellite images automatically. Nevertheless, it lacks the semantic information about the physical meaning of the different clusters. On the other hand, the second supervised layer contains the semantic knowledge of the vegetation patterns in a satellite image, and it can be trained with minimal training data. The proposed approach's outcome is a geolocation map for vegetation-related threats along power lines that provides updated situational awareness to vegetation management teams in electric utilities. The vegetation threat map is based on the density and proximity of vegetation encroachment in power lines' right-of-way. The proposed framework is implemented and validated in a vegetation management system for a power distribution company in the western part of Norway. The vegetation detection results from satellite images showed high matching with the available LiDAR survey data which has been used as the ground truth for the use case area.

## II. USE CASE AND DATA DESCRIPTION

The study has been performed in collaboration with a power distribution system operator (DSO) located in the western part of Norway. The study area is a 22kV sub-transmission network in a rural region that includes fields, sparse and dense forests, and water streams. Power lines' right-of-way is 20 meters on each side which forms a 40 meter corridor (see Fig. 1). Different datasets relative to the study area have been acquired as further described in the following.

### A. Satellite imagery

Two commercial high-resolution multispectral satellite images were used for the study area. The first is a Worldview-2 8-channel image provided by Maxar and acquired in May 2018. The second is a Pleiades-1 4-channel image provided by Airbus and acquired in September 2017, (see Fig. 2). Both images contain separate channels ranging from visible to near-infrared with a 0.5 meter/pixel spatial resolution.
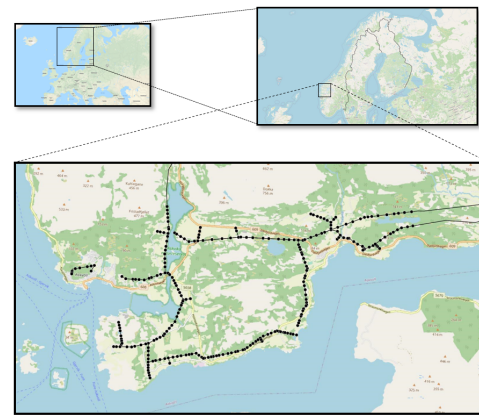


Fig. 1: Study area located on the western coast of Norway



WorldView-2: May 2018

(a)



Pleiades-1: Sept 2017

(b)

Fig. 2: RGB satellite image for the study area. The figure highlights the regional power line (red) and sub-regional power line (green).

### B. LiDAR point clouds

We also use LiDAR data in our study as a benchmark to validate our satellite-based vegetation detection. The available LiDAR data are grouped into different categories (vegetation, buildings, roads, stones, poles), see Fig. 3.

The heights of 22kV overhead lines in the study area in the western part of Norway are at a minimum of 7.8 meters with an average of 10.8 meters. Trees should have at least a 2.0 m distance to the conductors considering the catenary curvature of wires [28], [29]. In this study, we used LiDAR data provided by our electric utility partner as the ground truth for trees' location in the vicinity of power lines' right-of-way. However, the acquisition time for the available LiDAR data
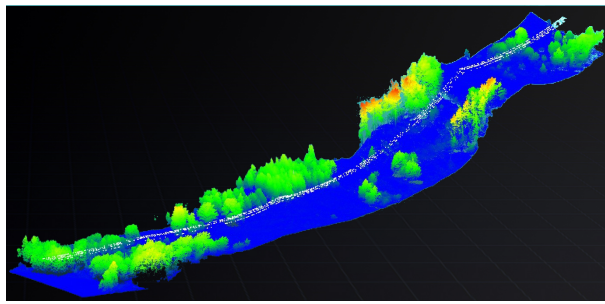
(a) 3D point clouds
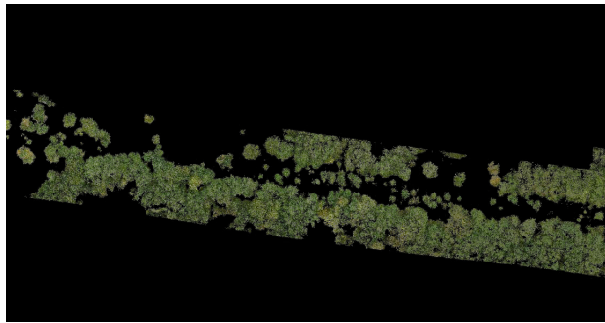


(b) 2D projection

Fig. 3: Examples of LiDAR data used in the study

(Sep 2019) is different from the satellite imagery data we have (Sep 2017 and May 2018) for the study area. To resolve the time difference among data from LiDAR and both satellite images, we assume that trees with a height higher than 2.5 meters in 2019, as observed in LiDAR data, are probably older than two years old based on growth rate of trees in this region. Therefore, the trees taller than 2.5 meters in 2019 are observable in satellite images from 2018 and 2017 at the same location along the power lines' ROW.

## III. Methodology

The satellite image is analyzed using the proposed algorithm, within a sliding window covering the power line right-of-way. In this study, we use windows of 40 by 40 meters (80 × 80 pixels). The algorithm sweeps the whole length of the electric power line in the study area.

Then, a segmentation map is created to show whether each pixel is a part of a tree or not. The proposed machine learning framework for satellite image processing is an ensemble of two different algorithms to enhance the overall performance. The first one is a supervised segmentation approach based on hand-crafted features, while the second one is a fully unsupervised algorithm developed for image segmentation tasks [30]. The output of the proposed machine learning framework is used for mapping vegetation risk along the power lines. Different blocks of our proposed framework, as shown in Fig. 4, are explained as follows.

### A. Data Pre-processing Block

To develop a learning algorithm for tree detection, we need to have labeled data for training and testing. Such labeled data are a collection of binary images (for example "1" for tree and

"0" for non-tree) that include the ground truth with the correct location of trees. In this paper, we create the first labeled dataset with the open-source raster graphics editor GIMP. A second set of labeled data is created automatically using LiDAR point-clouds for the same region. The 3D point clouds are projected in 2D at Nadir and converted into gray-scale. The resultant gray-scale binary image is smoothed through a dilation operation (using a $3 \times 3$ kernel).

The satellite image that we use is ortho-rectified and pan-sharpened [31]. Ortho-rectification enables the correction of potential defaults that exist due to satellite tilt or terrain distortions in cases where the satellite on-board sensor is not pointing directly at the Nadir direction. Pan-sharpening, on the other hand, increases the natively low-resolution parts of a multispectral image by combining them with the higher resolution panchromatic pixels. The image is in GeoTiff format, which includes geo-references for each pixel, allowing for proper location in real-world coordinates.

### B. Supervised Image Segmentation Block

Specific features are extracted at pixel-level, directly from the multi-channel images, and grouped into a vector associated with each pixel. A machine learning classifier [32] is then trained. In this paper, we use spectral features, texture features, and a Gaussian kernel as follows:

*1) Spectral Features:* First, we extract the pixel values coming from the different bands of a multispectral satellite image. Furthermore, we use the Normalized Difference Vegetation Index (NDVI) [33] to recognize vegetation. NDVI is a commonly used tool in remote sensing for vegetation detection and defined as:

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \qquad (1)$$

where $\rho_{red}$ and $\rho_{nir}$ stand for the spectral reflectance measurements acquired in the red (visible) and near-infrared regions, respectively.

*2) Texture Features:* Texture patterns are useful in identifying objects that may appear very similar to each other in an image from a color-based perspective (for example, trees and green fields). We convolute the gray-scale image obtained from the RGB components with a set of filter banks composed by Gabor filters to generate pixel responses at different scales and orientations [34].

As additional texture information, we also use the Gray Level Co-Occurrence Matrix (GLCM) [35]. From such a matrix, specific texture operators [36] can be extracted. Although a large number of operators exist, most of them are correlated, as explained in [37]. Therefore, we decided to use only contrast and correlation operators.

*3) Gaussian Kernels:* The next feature we use in this study is a Gaussian kernel. The Gaussian kernel (with variance equal to 1) is convolved with a gray-scale image derived from the RGB components. Such a kernel acts as a low-pass filter leading to a slightly blurred image. The Gaussian low pass filters are becoming more common in image processing to cancel the noise [38].
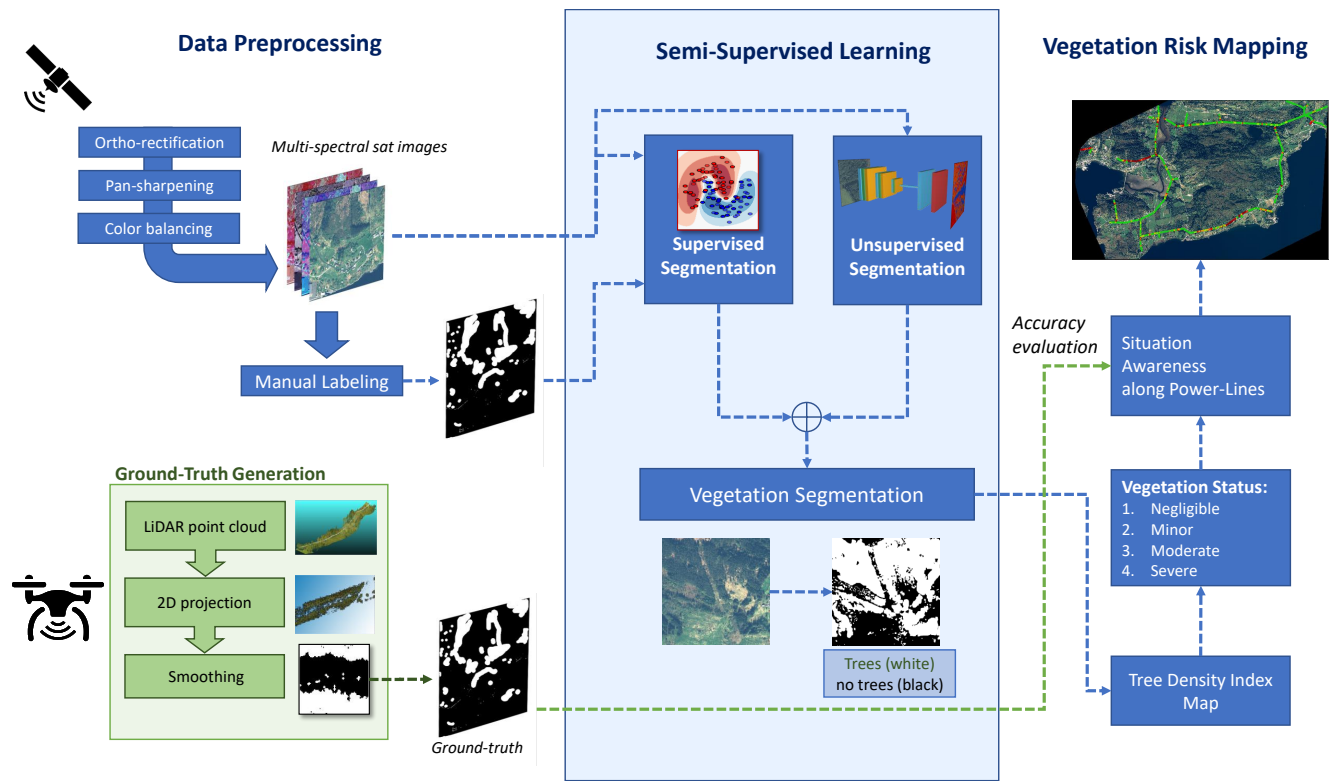
Fig. 4: Overview of the proposed vegetation monitoring framework

*4) Supervised Machine Learning Algorithm for Vegetation Segmentation:* The computed features values are stacked into a vector and we use the AdaBoost ensemble technique [39] to train a classifier. Such a classifier will assign a probability of being part of a tree, $P_{tree}$, to each pixel in the image. Finally, we use an energy minimization algorithm solved via graph cuts [40] to turn the probabilistic map into a binary segmentation map.

### C. Unsupervised Image Segmentation Block

The unsupervised segmentation block is composed of a fully convolutional neural network (FCN) [41] to extract features and a superpixel refinement process [42] for self-training of the model. Fig. 5 illustrates the architecture of this network.
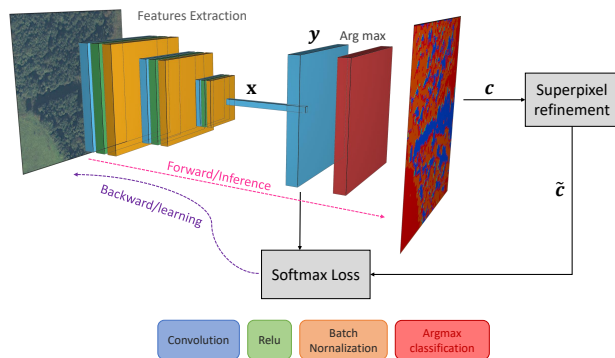


Fig. 5: Diagram of the unsupervised segmentation block

We compute the feature map $\mathbf{x}$ from the image $I$ through M (equal to 4 in our study) convolutional blocks consisting

each of a 2D convolution with a $3 \times 3$ kernel, a *relu* activation function, and a batch normalization step. Then, a response map is calculated through an additional convolution as $\mathbf{y} = W_c\mathbf{x} + b_c$ where $W_c, b_c$ are, respectively, the weights and biases of the last convolutional layer. Finally, we obtain the cluster label $c$ for each pixel by selecting the dimension along the vector $\mathbf{y}$ that has the maximum value.

In image segmentation, the clusters of image pixels should be spatially continuous. Therefore, we first extract $K$ fine superpixels from the image. Then, we force all the pixels in each superpixel to have the same cluster label $\tilde{c}$, defined as the most frequent cluster in each superpixel.

The self-training procedure is done solving two sub-problems alternately: a forward process of the network followed by the superpixel refinement and a back-propagation process based on stochastic gradient descent with a learning rate of 0.01. The loss function is calculated as the cross-entropy between the network response $\mathbf{y}$ and $\tilde{c}$. Algorithm 1 shows the pseudo-code of such an approach.

### D. Combination Strategy

The aforementioned fully unsupervised approach can segment the image into different clusters. A cluster is a group of pixels sharing common properties (color, texture pattern, etc). However, the algorithm lacks the semantic knowledge about the physical meaning of different clusters available in an image. On the other hand, the supervised model has been specifically trained to recognize trees. Therefore, we combine the output of the supervised block presented in subsection III-B with the output of the unsupervised block of subsection

**Algorithm 1** Unsupervised image segmentation

> **Input:** Image $I$
> **Output:** Segmented image $\mathcal{S}$
> $(W_m, b_m, W_c, b_c) \leftarrow$ InitializeWeights()
> $S_k \leftarrow$ GetSuperPixels($I$)
> **for** $t = 1$ to $N$ iterations **do**
>     $\mathbf{x} \leftarrow$ getFeatures($I; W_m, b_m)|_{m=1}^M$
>     $\mathbf{y} \leftarrow$ forwardStep($\mathbf{x}; W_c, b_c$)
>     $c \leftarrow \arg\max\{y\}$
>     **for** k=1 to K **do**
>         $c_{max} \leftarrow \arg\max c, \quad \forall c \in S_k$
>         $\tilde{c} \leftarrow c_{max}, \quad \forall pixel \in S_k$
>     **end for**
>     $\mathcal{L} \leftarrow CrossEntropyLoss(\mathbf{y}, \tilde{c})$
>     $(W_m, b_m, W_c, b_c) \leftarrow$ Update($\mathcal{L}$)
> **end for**

III-C. In this way it is possible to understand whether a cluster should be considered as "trees" or "non-trees". Algorithm 2 shows how this combination is performed.

**Algorithm 2** Combination of the two blocks

> **Input1:** Multi-class segmentation image from the unsupervised block $(\mathbf{U}) = \{U_0, U_1, U_2, \dots\}$
> **Input2:** Binary segmentation image from the supervised block $(\mathbf{S}) = \{S_0 = \text{trees}, S_1 = \text{no-trees}\}$
> **for** each cluster $U_i$ in $(\mathbf{U})$ **do**
>     Check in which class of $(\mathbf{S})$ the pixels of $U_i$ are mapped into
>     **if** most of the pixels are mapped into $S_0 = 0$ **then**
>         Assign the pixels of $U_i$ to $S_0$
>     **end if**
>     **if** most of the pixels are mapped into $S_1$ **then**
>         Assign the pixels of $U_i$ to $S_1 = 1$
>     **end if**
> **end for**

## IV. RESULTS AND DISCUSSIONS

To validate the performance of the proposed framework for vegetation detection, we implement it on a 22kV sub-transmission power line in the western part of Norway with 26 km of lines. Two high-resolution satellite images have been used, as explained in subsection II-A. The available LiDAR survey for the same part of lines has been used as the ground truth to cross-validate our satellite-based solution's outcomes.

To start the validation, we test the supervised block to compute the segmentation output accuracy by applying different combinations of features. We use the manually-annotated dataset for the satellite imageries. The training dataset consists of five manually labeled 800x800 pixels images, where a subset of 1.500.000 pixels has been selected to have balanced classes. The validation dataset is made of ten 400x400 pixels images.

We found out that the GLCM texture operator performs better than the Gabor filter. Surprisingly, adding the NDVI lowers the overall accuracy. NDVI is an indicator of the chlorophyll richness, so it can be used to detect vegetation. However, it fails to sufficiently distinguish between trees and grassy fields. Furthermore, trees with small canopies or otherwise sparse foliage might not be adequately detected using NDVI. However, NDVI remains an important vegetation

detection index that may still be helpful in other scenarios with different datasets.

The machine learning algorithm output is a black and white segmented image that shows tree and non-tree in each part of the line, (see the blue-colored zone in the middle of the Fig. 4). For the sake of visualization, Fig. 6 shows a comparison between the manually labeled satellite images and the classifier's output for three image samples. A comparison between the proposed approach's output and LiDAR's ground truth is shown in Fig. 7.
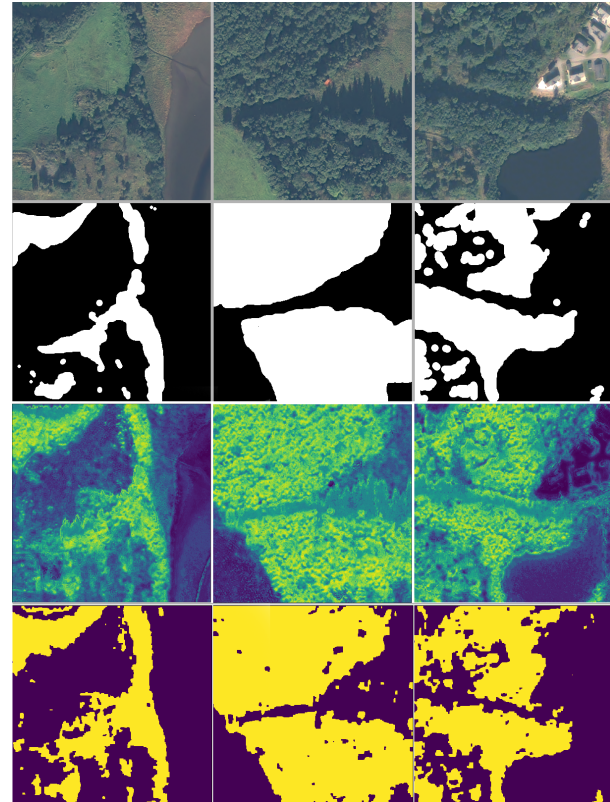


Fig. 6: From top to bottom: RGB image *(first row)*, ground truth provided by manual labeling *(second row)*, probabilistic map showing $P_{tree}$ for each pixel *(third row)*, and segmentation output *(last row)*

.

We create an easy-to-understand metric, called *Tree Density Index* (TDI), to present and visualize our vegetation detection algorithm's outcome for grid maintenance teams. The proposed metric can be used in mapping vegetation encroachment to power lines right-of-way. The TDI quantifies the presence of trees near power lines, especially the trees that encroach into the lines' ROW.

We multiply the segmented images $M$ with a Gaussian kernel $G$ as a weighting function within the window $W$, as described in Eq. (2). The weighting function's choice (TDI) is based on the distance of trees to power lines, since trees near power lines pose more risk.

$$TDI = \int_W M \otimes G \, dw \in (0, 1) \qquad (2)$$

The TDI values, in the range of [0,1], are divided into different levels of vegetation status using the following criteria.
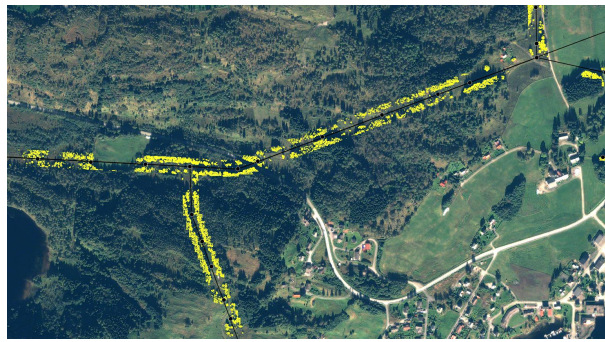
(a) Tree coverage scanned by LiDAR
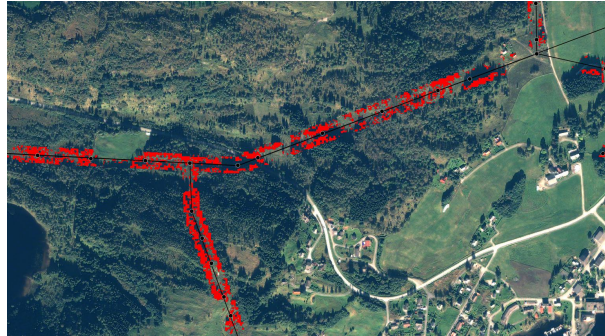


(b) Tree coverage detected by the algorithm

Fig. 7: Example of trees detected along the power grid by (a) LiDAR and (b) the proposed classifier



0 < TDI < 0.2    0.2 < TDI < 0.4    0.4 < TDI < 0.6    0.6 < TDI < 0.8    0.8 < TDI < 1

Fig. 8: Heat-map showing areas where there is more vegetation around power lines



(a) Pleiades-1 satellite image    (b) WorldView-2 satellite image

Fig. 9: Confusion matrix from the levelized values of the TDI along the line using the Pleiades-1 and WorldView-2 satellite image

$$
\begin{cases}
\text{Level 0 (Negligible):} & TDI \leq 0.2 \\
\text{Level 1 (Minor):} & 0.2 < TDI \leq 0.4 \\
\text{Level 2 (Moderate):} & 0.4 < TDI \leq 0.8 \\
\text{Level 3 (Severe):} & TDI > 0.8
\end{cases}
\tag{3}
$$

Using the Tree-Density Index from Eq. (2), it is possible to create a heat map showing vegetation density and proximity levels along the power grid. We calculate TDI values for the entirety of the power lines in our study area using the two satellite images and the LiDAR ground truth. Fig. 8 shows an example of the resultant heat map for the study area. The vegetation density heat map shows that most of the line sections are safe (green colored and TDI < 0.2). It also shows that our partner electric utility does not need to make an immediate tree trimming action in those areas.

A confusion matrix is then used to show the comparison results for detected trees' location in satellite images and LiDAR data. Fig. 9a and 9b show the confusion matrices for Pleiades-1 and WorldView-2 imagery respectively. The upper-triangular part of each confusion matrix corresponds to locations where the predicted TDI is higher than the true value; it means that we are overestimating the vegetation density. Symmetrically, the lower-triangular part corresponds to areas where the predicted TDI is lower than the real value; in other words, it shows that we are underestimating the vegetation density.

We make the assumption that all locations with TDI values higher than 0.4 can cause a vegetation related threat to power lines. From Eq. (3), va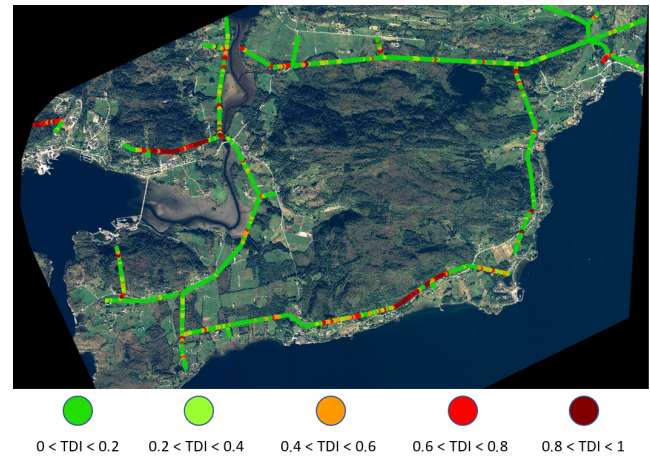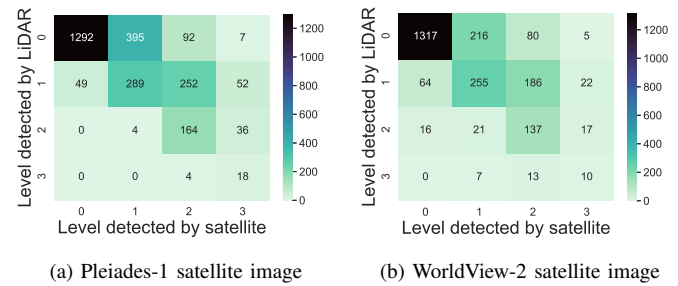lues higher than 0.4 cover areas with moderate to high vegetation densities, and they need to be monitored carefully for possible tree trimming and cutting actions. In this way, we can compute how well the algorithm detects a vegetation threat. For example, using Pleiades-1 satellite image, the algorithm detects non-threat vegetation zones correctly by 98.2% and it detects the threat vegetation zones by 84.6%.

From Fig. 9, we also see that there are very few cases in which the algorithm underestimates the vegetation level. For example, using Pleiades-1 there are zero cases in which a location detected as *negligible* (level 0) is *severe* (level 3) in reality.

In practice, this means that our algorithm has high confidence in detecting areas with low vegetation density. In other words, the vegetation management team can avoid inspecting areas that the algorithm points to as "safe" (green) or without risky vegetation. Consequently, it brings the vegetation management teams' attention on areas with high vegetation density or with risky vegetation (red). It reduces the power line inspection time and cost accordingly.

## V. PRACTICAL CONSIDERATIONS FOR ELECTRIC UTILITIES

The combination of satellite-based data and artificial intelligence gives electric utilities a unique opportunity to modernize tasks that incorporated repetitive observation and inspection,

especially over large areas. This paper introduces a platform for monitoring vegetation encroachment into power lines' right-of-way using high-resolution satellite imagery.

## A. On the classic close-up asset monitoring and satellite images

It is worth mentioning that the need for close-up inspection of power line components using ground patrol, helicopters, or drones to check the asset's mechanical and structural health, in addition to vegetation monitoring, remains the same. In reality, classic close-up visual inspections for the whole grid are performed once within a long period (up to 10 years) due to their high costs and time constraints. This infrequent inspection over a long span of time can increase the probability of failures and outages specifically for vegetation-related events considering vegetation growth rates. Therefore, vegetation-related inspections need to be executed more often and in periods between classic close-up asset inspections. Our proposed satellite-based vegetation monitoring approach complements the legacy asset management practice by providing low-cost and frequent situational awareness for vegetation management teams.

Consequently, vegetation management can be changed from traditional time-based ROW inspection (periodic) to risk-based ROW inspection by improving scale, frequency, and cost-efficiency in ROW inspection.

## B. On the cost-effectiveness of satellite images for vegetation monitoring

Satellite imagery data is typically more cost-effective than other image capture methods such as helicopter and drone, especially as the inspection area increases [11], [12]. Commercial satellite providers can offer high-resolution images with a high revisiting time covering most of the world. A survey performed in 2015 [43] showed that satellite imagery for a specific region was up to 60 % cheaper than using drone images.

Our study was performed under the GridEyeS project for using satellite imagery for power system operation supported by the European Space Agency [44]. We have surveyed 15 electric utilities in North America and Europe regarding their typical practices and cost of vegetation monitoring during our study. The line inspection cost using helicopters, light airplanes, or drones varies from 60 to 1300 Euros per km of the power line. The higher range of inspection costs belongs to LiDAR scanning technologies. The high cost of power lines' health condition monitoring (including vegetation encroachment monitoring) and the vast size of service territories force utilities to often cover the whole service area with a long periodicity (typically 2 to 10 years) [5]. This leads to suboptimal revisiting frequency for each section of the line.

The use of high-resolution satellite imagery for vegetation-related inspection costs is generally below 15 Euros, depending on commercial providers. This makes satellite-based solutions economically attractive. The power line in our selected study area is approximately 26 kilometers long. In total, 20 kilometers of the line are in a normal condition regarding

vegetation encroachment into power lines' ROW. Using the proposed platform, the vegetation management team has an option to only focus on the 6 km of the line with a high vegetation level, identifiable with the red color in Fig. 8.

The average cost of line inspection for our partner utility using a helicopter or a drone is 1200 Euros per km, which means an overall 26 x 1200 = 31200 Euros in the inspection cost. Limiting vegetation-related inspections to only 6 km of the red zone leads to inspection cost reduction depending on the distribution of red spots across the entire line and the number of helicopter or ground crew maneuvers to cover those spots. This example aims to provide an idea of saving potential for vegetation monitoring rather than a detailed cost-benefit analysis.

## C. On the selecting the appropriate satellite images for vegetation monitoring

From a practical point of view, remote sensing applications' accuracy, including vegetation detection, is highly influenced by different aspects.

- A major factor in the quality of vegetation detection is the resolution of the satellite images. In this study, we used 50 cm resolution satellite images, which are one of the best resolution available from commercial providers. For example, there are no-cost to low-cost satellite images with 10 meters resolution from Sentinel satellites under the Copernicus program provided by the European Space Agency [45]. However, such low-resolution images can not provide the level of details needed to detect and measure vegetation's density in the vicinity of power lines. Moreover, the high-resolution satellite images also provide insights on vegetation type, growth rate, vegetation health, environmental impacts, and the quality of trimming activities by using data from multiple satellite images of the same area over time. Figure 10 shows how the different resolution of some available products can dramatically affect the detection quality.
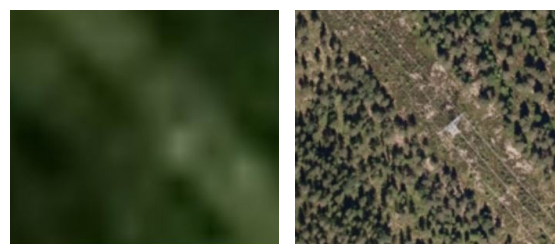


(a) 10 meter resolution satellite image    (b) 0.5 meter resolution satellite image

Fig. 10: Visual comparison of the same area from two different popular satellite providers

- Another important aspect is the acquisition date. As shown in the Results section, Pleiades-1 provided slightly better accuracy. A possible explanation is that in early May, trees are not yet completely developed, especially in Norway. Therefore, the algorithm has more difficulties in precisely detecting trees, particularly deciduous ones.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3059307, IEEE Transactions on Power Delivery

8

Fig. 11 shows a comparison between two image samples from different seasons. Note that the foliage of some trees are not as fully developed in May compared to the canopy cover in September. It is worth mentioning that there can be periods more suitable for vegetation monitoring. Power companies should prioritize monitoring in specific months rather than in other months (in summer, for example) to generate the best results, taking geographical location and growing seasons into consideration.



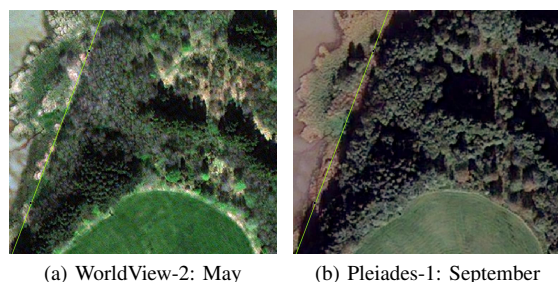(a) WorldView-2: May      (b) Pleiades-1: September

Fig. 11: Visual comparison between two images before and after summer. Note how the canopy of some trees are not fully developed yet in May compared to September

- Cloud coverage is also another factor that affects all remote sensing applications which are based on optical imagery. Prior to our analytics, we tried to pick proper satellite images that have less cloud coverage over the target area. Nowadays, commercial satellite image providers are launching more satellites to orbit using private companies such as Space-X. More satellites in the space means more frequent revisit for any location in the world. Consequently, the recent increase in satellite revisit frequency will make it easier to acquire a cloudless image for a specific area with a short waiting time.

## VI. Conclusions

This paper presents a framework to monitor vegetation along power lines using high-resolution satellite images and machine learning. Satellite imagery data introduce a new paradigm for power transmission and distribution companies with the potential to reduce the time and cost of ground inspections. We propose a semi-supervised approach that combines a supervised classifier with a deep learning-based unsupervised architecture for image segmentation. This enables the detection of vegetation close to power lines and thus pose a risk to power line infrastructure. The proposed framework has the potential to aid operators of power-line infrastructure with vegetation management. We validated the image segmentation approach for a power grid in western Norway using airborne LiDAR. Initial results indicate that this approach can correctly identify vegetation risk areas with 84% accuracy for this particular area. Areas of no risk are identified correctly in 92% of cases. These initial results demonstrate the potential promise of this satellite-based framework. The future work is toward further improving the detection accuracy.

## VII. Acknowledgments

## References

[1] H. Gugel, S. Ekisheva, M. Lauby, and F. Tafreshi. Vegetation-related outages on transmission lines in north america. In *2018 IEEE Power Energy Society General Meeting (PESGM)*, pages 1–5, 2018.

[2] Yoshitaka Kumagai, John C. Bliss, Steven E. Daniels, and Matthew S. Carroll. Research on causal attribution of wildfire: An exploratory multiple-methods approach. *Society & Natural Resources*, 17(2):113–127, 2004.

[3] P. Eavis and I. Penn. California says pg&e power lines caused camp fire that killed 85. https://www.nytimes.com/2019/05/15/business/pge-fire.html?action=click&module=RelatedLinks&pgtype=Article. 2019-05-15.

[4] J. Ahmad, A. Malik, L. Xia, and Nadia Ashikin. Vegetation encroachment monitoring for transmission lines right-of-ways: A survey. *Electric Power Systems Research*, 95:339–352, 2013.

[5] B. Sacks. Ai improves resilience of vegetation management. https://www.tdworld.com/vegetation-management/article/21131079/ai-improves-resilience-of-vegetation-management. 2020-05-11.

[6] J. Weeks. U.s. electrical grid undergoes massive transition to connect to renewables. https://www.scientificamerican.com/article/what-is-the-smart-grid/. 2010-04-28.

[7] Y. Xie, Zongyao Sha, and Mei Yu. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology*, 1:9–23, 2008.

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

[9] R. A. McLaughlin. Extracting transmission lines from airborne lidar data. *IEEE Geoscience and Remote Sensing Letters*, 3(2):222–226, 2006.

[10] Lake Singh, William Whittecar, Marc DiPrinzio, Jonathan Herman, Matthew Ferringer, and Patrick Reed. Low cost satellite constellations for nearly continuous global coverage. *Nature Communications*, 11, 01 2020.

[11] Ucs satellite database. https://www.ucsusa.org/nuclear-weapons/space-weapons/satellite-database. 2020-08-01.

[12] H. Jones. The recent large reduction in space launch cost. 2018.

[13] Y. Kobayashi, G. G. Karady, G. T. Heydt, and R. G. Olsen. The utilization of satellite images to identify trees endangering transmission lines. *IEEE Transactions on Power Delivery*, 24(3):1703–1709, 2009.

[14] M. Moeller. Monitoring powerline corridors with stereo satellite imagery. 2006.

[15] C. Xiao, R. Qin, X. Huang, and J. Li. Individual tree detection from multi-view satellite images. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 3967–3970, 2018.

[16] Grigorijs Goldbergs, S. Maier, S. Levick, and Andrew Edwards. Limitations of high resolution satellite stereo imagery for estimating canopy height in australian tropical savannas. *Int. J. Appl. Earth Obs. Geoinformation*, 75:83–95, 2019.

[17] S. J. Mills, M. P. Gerardo Castro, Z. Li, J. Cai, R. Hayward, L. Mejias, and R. A. Walker. Evaluation of aerial remote sensing techniques for vegetation management in power-line corridors. *IEEE Transactions on Geoscience and Remote Sensing*, 48(9):3379–3390, 2010.

[18] X. Huang, C. Shi, and S. C. Liew. Tree crown detection and delineation using optical satellite imagery. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 2944–2947, 2018.

[19] Z. Pan. Urban vegetation type analysis method based on high resolution satellite images. In *2016 International Conference on Smart City and Systems Engineering (ICSCSE)*, pages 613–616, 2016.

[20] L. Häme, J. Norppa, P. Salovaara, and J. Pylvänäinen. Power line monitoring using optical satellite data. In *CIRED Workshop 2016*, pages 1–4, 2016.

[21] Yutaka Kokubu, Seiichi Hara, and Akira Tani. Mapping seasonal tree canopy cover and leaf area using worldview-2/3 satellite imagery: A megacity-scale case study in tokyo urban area. *Remote Sensing*, 12(9):1505, May 2020.

[22] Pedro Marques, L. Pádua, Telmo Adão, Jonás Hruska, E. Peres, A. M. Sousa, and J. Sousa. Uav-based automatic detection and monitoring of chestnut trees. *Remote. Sens.*, 11:855, 2019.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3059307, IEEE Transactions on Power Delivery

9

[23] S. Li, W. Song, Leyuan Fang, Yushi Chen, P. Ghamisi, and J. Benediktsson. Deep learning for hyperspectral image classification: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57:6690–6709, 2019.

[24] Ronald Kemker, C. Salvaggio, and Christopher Kanan. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. *Isprs Journal of Photogrammetry and Remote Sensing*, 145:60–77, 2018.

[25] B. Ayhan and C. Kwan. Tree, shrub, and grass classification using only rgb images. *Remote. Sens.*, 12:1333, 2020.

[26] F. Wagner, A. Sánchez, Yuliya Tarabalka, R. G. Lotte, Matheus Pinheiro Ferreira, M. P. Aidar, E. Gloor, O. Phillips, and L. Aragão. Using the u-net convolutional network to map forest types and disturbance in the atlantic rainforest with very high resolution images. *Remote Sensing in Ecology and Conservation*, 5:360–375, 2019.

[27] Sherrie Wang, William Chen, Sang Xie, George Azzari, and David Lobell. Weakly supervised deep learning for segmentation of remote sensing imagery. *Remote Sensing*, 12:207, 01 2020.

[28] https://www.lysenett.no/byggeoggrave/grave-og-arbeide-nar/trefelling-og-beskjaring/.

[29] Knut Ove Hillestad. Kraft, ledning og landskap. http://publikasjoner.nve.no/kraftogmiljoe/kraftogmiljoe_08.pdf.

[30] A. Kanezaki. Unsupervised image segmentation by backpropagation. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1543–1547, 2018.

[31] Snehmani, Akshay Gore, Ashwagosh Ganju, S. Kumar, P. K. Srivastava, and P HariRamR. A comparative analysis of pansharpening techniques on quickbird and worldview-3 images. *Geocarto International*, 32:1268 – 1284, 2017.

[32] Jerome Friedman Trevor Hastie, Robert Tibshirani. *The Elements of Statistical Learning*. Springer, 2009.

[33] https://en.wikipedia.org/wiki/Normalized_difference_vegetation_index.

[34] U. Marmol. Use of gabor filters for texture classification of airborne images and lidar data. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 22, 2011.

[35] Mryka Hall-Beyer. Glcm texture: A tutorial v. 3.0 march 2017, 03 2017.

[36] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, 1973.

[37] Mryka Hall-Beyer. Practical guidelines for choosing glcm textures to use in landscape classification tasks over a range of moderate spatial scales. *International Journal of Remote Sensing*, 38:1312 – 1338, 2017.

[38] Pascal Getreuer. A survey of gaussian convolution algorithms. *Image Process. Line*, 3:286–310, 2013.

[39] R. E. Schapire and Y. Freund. *Boosting: Foundations and Algorithms*. MIT Press, 2012.

[40] Yuri Boykov and Olga Veksler. Graph cuts in vision and graphics: Theories and applications. In *Handbook of Mathematical Models in Computer Vision*, 2006.

[41] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440, 2015.

[42] Murong Wang, Xiabi Liu, Yixuan Gao, Xiao Ma, and N. Q. Soomro. Superpixel segmentation: A benchmark. *Signal Process. Image Commun.*, 56:28–39, 2017.

[43] Eyes in the sky: Satellite remote sensing and data analytics for electric utilities. https://www.epri.com/research/products/000000003002017281, 2019-09-10.

[44] https://business.esa.int/projects/grideyes.

[45] https://www.esa.int/Applications/Observing_the_Earth/Copernicus.

**Michele Gazzea, IEEE Student Member** Michele Gazzea received his Bachelor's degree in Information Engineering and his Master's degree in Automation and Control Engineering, both from the University of Padova (Italy). He worked for one year in Cielle.srl in Treviso (Italy) as an R&D engineer studying and designing diagnostic techniques on milling and engraving CNC machines. He worked as a researcher for Electrolux to perform model-based analysis of power consumption in washing machines until he started his new occupation as a Ph.D. student in Western Norway University of Applied Sciences in Bergen (Norway). His research interests are in data analytics, machine learning, computer vision, and remote sensing applications.
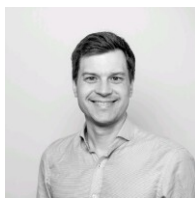
**Michael Pacevicius** Michael Pacevicius is an industrial PhD candidate in the RAMS group at NTNU, Norway and works as a researcher in eSmart Systems, a company delivering high-tech IT solutions for power grid related companies. His research activities focus on the development and implementation of dynamic risk analysis methods for large-scale interconnected power systems. He has a MSc. in Operational safety, Risks and Environment from the Université de Technologie de Troyes (UTT) in France and a MSc. in Economics and Business Administration from the Technische Universität Braunschweig (TUBS) in Germany. He worked as a project coordinator and analyst in the Big Data business development department of SAP in Munich, Germany, before joining eSmart Systems back in 2017.
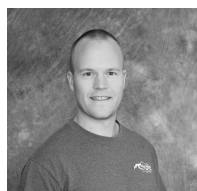
**Dyre Oliver Dammann** Dyre Oliver Dammann received the Ph.D. degree in geophysics in 2017 from the University of Alaska Fairbanks, Fairbanks, AK, USA. From 2017-2019 he was a Postdoctoral Researcher in the Department of Earth, Space, and Environment, Chalmers University of Technology, Gothenburg, Sweden. From 2019-2020 he was a Senior Scientist with StormGeo, Bergen, Norway. He is currently a Research Professional at the University of Alaska Fairbanks. His research interests include coastal sea ice properties and processes centered around sea ice system services and emerging needs of Arctic stakeholders. His research is particularly focused on expanding monitoring tools using synthetic aperture radar in support of ice travel and on-ice operations of landfast sea ice.

**Alla Sapronova** Alla Sapronova, Ph.D., has experience in artificial intelligence, machine learning, and data analysis. She graduated from Moscow State University, Russia, in 2004 with Ph.D. in Physics and Mathematics with Honored President's Stipend. After completing the postdoctoral program at the University of Bergen, Norway, in 2008, she started to work with artificial intelligence as a Senior Researcher at the Bergen Center for Computational Science affiliated with the University of Bergen, Norway. From 2014 Sapronova was appointed as a Head of Data Science for the Big Data Analysis Center at the University of Bergen, with responsibility ranging from project management and data science team leading to external project building and customer relationship. From 2019 she is a Lead Data Scientist at the StormGeo company. During her career as Data Scientist, she completed projects within artificial intelligence and machine learning for various industries and public sectors, ranging from medical hospitals and fisheries to oil drilling and insurance companies. Sapronova has over 20 scientific publications for the past 10 years.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3059307, IEEE Transactions on Power Delivery

10

**Torleif Markussen Lunde** Torleif Markussen Lunde is Managing Director of Innovation Centre for Health at the University of Bergen. He has worked in the intersection between industry and research for the last 10 years years and has experience from a leadership positions related to software development, strategy, research, and transformation. Torleif is passionate about technology and people that delivers real impact and solve real-world problems. Holds a PhD in mathematical modelling & big data, climate change, and malaria.

**Reza Arghandeh, IEEE Senior Member** Dr. Reza Arghandeh is a Full Professor in the Department of Computing, Mathematics, and Physics and Department of Electrical Engineering at the Western Norway University of Applied Sciences (HVL), Norway. He is the Director of Collaborative Intelligent Infrastructure Lab (CI2). He is also a lead data scientist with StormGeo AS. He has been an Assistant Professor in ECE Dept, FSU, USA 2015-2018, and a postdoctoral scholar at EECS Dept, University of California, Berkeley 2013-2015. He was a power system software designer at Electrical Distribution Design Inc. in Virginia, USA, from 2011 to 2013. His research has been supported by U.S. National Science Foundation, U.S. Department of Energy, the European Space Agency, and the European Commission.