Plant Species-Specific Anomaly Detection Based on Electrophysiological Signals

Andy Desman Lo, Elvin Nur Furqon, Junaidul Islam, Isack Farady,
Kahlil Muchtar, Ronnie Concepcion II and Chih-Yang Lin*
National Central University, Taoyuan, Taiwan
E-mail: andydesmanlo@gmail.com
National Central University, Taoyuan, Taiwan
E-mail: 113383602@cc.ncu.edu.tw
National Central University, Taoyuan, Taiwan
E-mail: imbj92@gmail.com
Yuan Ze University, Taoyuan, Taiwan
E-mail: isackf@saturn.yzu.edu.tw
Universitas Syiah Kuala, Banda Aceh, Indonesia
E-mail: kahlil@usk.ac.id
De La Salle University, Manila, Philippine
E-mail: ronnie.concepcion@dlsu.edu.ph
*National Central University, Taoyuan, Taiwan
E-mail: andrewlin@ncu.edu.tw

***Abstract*—Early detection of plant responses to stress and pathogen stimuli is essential for maintaining crop health and optimizing yield. Electrophysiological signals provide a non-invasive, real-time method for monitoring plant response. This study explores the application of a one-dimensional autoencoder for detecting abnormal patterns in electrophysiological signals recorded from plant branches. The autoencoder achieves this by learning to compress the input signals and reconstruct them from the compressed signal. Signals that show high reconstruction errors are considered to significantly differ from the learned normal patterns and are flagged as potential anomalies. Signals from multiple plant species such as *Araucaria, Cyperus, Hedera, Mentha, Ocimum, Plectranthus, Ruta, Salvia, Solanum* were preprocessed digitally using detrending, temporal shifting, and normalization techniques. Model performance was evaluated using Mean Absolute Error (MAE) and Intersection over Union (IoU) metrics. Results indicate that reliable anomaly detection was achieved for four species *Ocimum* (IoU: 0.615), *Plectranthus* (0.701), *Ruta* (0.631), and *Solanum* (0.670), while other species exhibited lower IoU scores, likely due to weaker or less distinguishable electrophysiological responses. The result also shows that while autoencoder approach can detect abnormalities in some plant species, their performance is highly correlated on species-specific signal pattern.**

Keywords*—bioelectrical, autoencoder, anomaly detection, signal processing, plant monitoring

I. INTRODUCTION

Electrophysiological signals such as action potentials, variation potentials, and related voltage changes are some of the earliest responses that plants produce when facing stress conditions like wounding, drought, salinity, or pathogen attack. These signals travel through the plant and help activate its defense mechanisms [1]. These signals happen before any visible symptoms shows such as wilting, discoloration, curling. so they can be used as an early warning system, allowing farmers to respond faster and protect crop yields. With the development of plant-friendly sensors and low-power recording devices, it is now possible to monitor these signals continuously in the field without harming the plants, leading to large volumes of useful data [2]. The main challenge lies in turning signal data into reliable indicators of plant health. The information we get from signal such as amplitude, shape, and baseline drift can vary between species, across growth stages, and even within the same plant. In addition, the lack of open-source and well-labeled datasets makes it hard to develop a supervised learning models.

In order to address the current challenge regarding the dataset and information from signal, one popular approach is to use reconstruction-based autoencoders, which learn to compress and rebuild normal signals. If a signal is significantly different from what the model has learned, it will show a high reconstruction error and can be marked as abnormal [3], [4]. Another method is clustering, such as K-means clustering, which defines regions of normal signals in the feature space and identifies points far from

979-8-3315-7206-8/25/$31.00 ©2025 IEEE

these clusters as outliers [5], [6]. Both approaches have shown reliable results in industrial and medical applications [7], [8]. However, their performance in plant electrophysiology is still unclear, especially since there is no established multi-species benchmark dataset.

The primary contribution of this research is the application of an autoencoder model, combined with detrending and normalization preprocessing methods, to detect plant anomaly responses to external stimuli demonstrates the potential of using electrophysiological signals based on the open dataset published by Madariaga et al. [9] for early warning applications.

II. RELATED WORKS

Many studies have shown that plants respond to stress through changes in their electrical signals [10], [11]. Research has demonstrated that different types of stress produce distinct waveforms. Mudrilov et al. [12] found that heat, burning, and mechanical damage each trigger unique electrical pulses. These differences in pulse height and recovery time act like fingerprints that reflect the specific condition affecting the plant.

Recent progress in sensor design and signal analysis is helping move electrophysiological monitoring closer to real-world use. Wen et al. [13] inserted flexible fiber-based electrochemical transistors into the xylem and discovered that hydrogen peroxide bursts travel with variation potentials during heat and light shocks. These bursts and electrical signals strengthened each other, showing how chemical and electrical responses are closely connected during long-distance signal transmission.

Zhou et al. [14] developed a microneedle electrode array that attaches to the surface of a leaf. They combined it with an on-board XGBoost model to detect in real time whether the plant is under drought, salt stress, or mechanical damage. In another study, Furch et al. [15] showed that when a plant detects a bacterial peptide (flg22), it sends out electrical signals that move through the vascular system, close the sieve tubes, and activate defense responses. This proves that electrical signals also carry information about microbial threats.

These findings shows that plants generate unique electrical signatures when exposed to different types of stress. By leveraging a flexible sensors and machine learning techniques, we can analyze these signals to determine the plant's condition. This enables early detection of stress, which helps save resources and improve crop productivity.

III. METHODS

To detect abnormalities in plant electrophysiological signals, we combined signal preprocessing with a semi-supervised anomaly detection model (Fig. 1). The raw signals were first detrended to remove baseline drift, then augmented to increase data diversity, and finally normalized to maintain a consistent scale across all recordings. A one-dimensional autoencoder trained only on normal data, using a sliding window approach to divide signals into smaller sections. During evaluation, the model flagged any input with a high reconstruction error that exceeded the predefined threshold as a potential anomaly.

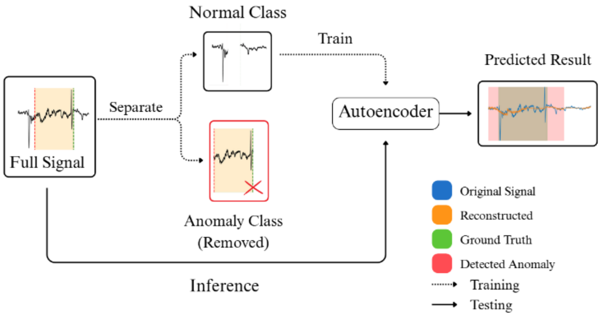

Fig. 1. Step-by-step process flow in detecting anomaly.

A. *Electrophysiological Signal Dataset*

Electrophysiological signals were recorded from sixteen different plant species by wrapping an electrode around the plant branch with conductive gel applied at the contact point. The electrode was connected to a Plant SpikerBox amplifier for signal acquisition. Each recording is annotated with markers that indicate the start (marker ID 1) and end (marker ID 2) of external stimuli [9]. For this study, only those species with recordings that included both markers, ensuring complete information about stimulus events are selected. During training, we used only the segments labeled as "*normal*," which are located outside the stimulus period. For evaluation, both normal and stimulus segments were included.

Table 1 summarizes the species included in this study, the number of recordings per species, and the reasons for excluding certain species. The reason some species were excluded from this study is primarily due to incomplete labeled data, such as missing end markers (marker id 2) for stimulus events or a limited number of recordings (only one or two available samples).

TABLE I. DATASET INFORMATION

Plant Species	**Usability**	**Total Signal**	**Remarks**
Araucaria	Yes	7 signals	-
Cyperus	Yes	8 signals	-
Dionaea	No	2 signals	Insufficient data
Drosera	No	-	Incomplete marker file
Hedera	Yes	9 signals	-
Melissa	Yes	15 signals	-
Mentha	Yes	16 signals	-
Mimosa	No	2 signals	Insufficient data
Ocimum	Yes	38 signals	-
Origanum	No	1 signal	Insufficient data
Peperomia	No	-	Incomplete marker file
Plectranthus	Yes	23 signals	-

Polypodiopsida	No	2 signals	Insufficient data
Ruta	Yes	36 signals	-
Salvia	Yes	30 signals	-
Solanum	Yes	17 signals	-

B. *Signal Pre-Processing*

To improve signal quality and increase training data, all raw signals were preprocessed using three steps:

- Detrending

Generally, detrending is used to remove slow changes or drifts baseline in the signal, which can be caused by sensor sensitivity or gradual environmental changes [16], [17]. Since the stimulus response in plant signals are usually quick and volatile, therefore, detrending helps us focus on these real and fast events rather than the slow background shifts.

The progress begins by identifying a linear trend in each signal through fitting a linear regression line (Fig. 2), using the least squares to estimate the slope (a), intercept (b) and time index (t).

$$x[t] \approx at + b \tag{1}$$

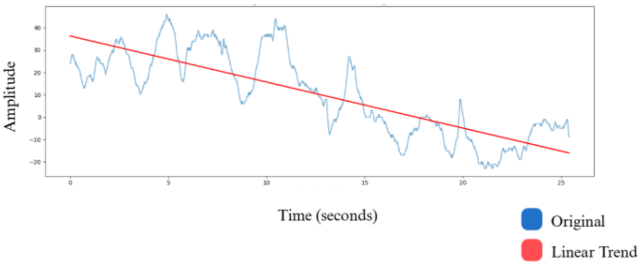

Fig. 2. Linear Regression Trend.

The detrended signal was then obtained by subtracting this linear trend (Fig. 3).

$$X_{Detrended}[t] = x[t] - (at + b) \tag{2}$$

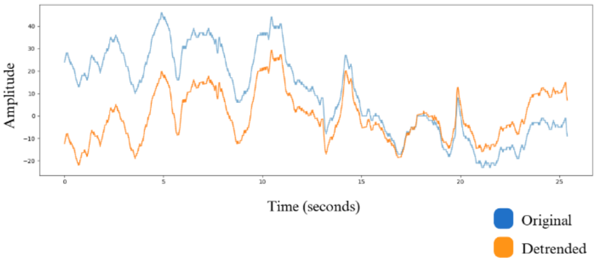

Fig. 3. Original and Detrended Signal Comparison.

- Augmentation

Due to the limited size of the signal data available for each species, we expanded the dataset by applying shift augmentation, in which the original signals are shifted along the temporal axis to generate additional samples., where the shift value s is randomly selected with a maximum offset of 1 s to the left or right This step increases the number of training samples from each original recording, which is especially valuable when available data is scarce.

$$X_{Shifted}[t] = X_{Detrended}[(t + s)] \tag{3}$$

- Normalization

Finally, all signals (original and augmented) were normalized to zero mean and unit variance, where μ and σ denote the mean and standard deviation of each shifted signal, respectively. This normalization ensures a consistent scale across all signals, enabling fair comparison and focusing analysis on the shift patterns rather than differences in raw amplitude.

$$X_{Norm}[t] = \frac{X_{Shifted}\,[t] - \mu}{\sigma} \tag{4}$$

These preprocessing steps remove unwanted baseline trends, expand the training set, and ensure that all signals are on a consistent scale.

C. *Autoencoder Modeling*

Autoencoder is a method designed for unsupervised representation learning and data reconstruction. It consists of an encoder and decoder. The goal of the encoder was to reduce the input data into a lower-dimensional latent space, and the decoder attempted to reconstruct the original input from this reduced representation as closely as possible, thereby minimizing the loss between the original and reconstructed input. During testing, the autoencoder reconstructed data that it had seen during training. When it encountered inputs that differed significantly from the training data, it produced high reconstruction errors. By leveraging this behavior, we detected anomalies by identifying input segments where the reconstruction errors exceeded the predefined threshold obtained during training [18].

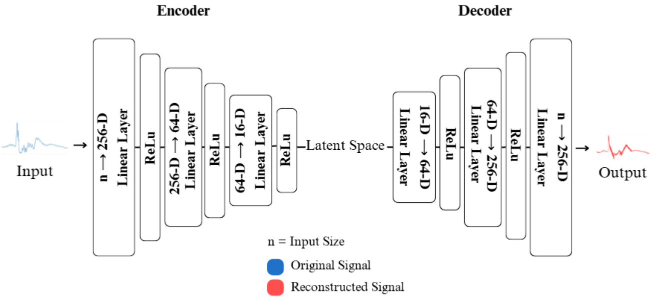

Fig. 4. Autoencoder network structure.

D. *Evaluation Metrics*

To evaluate the performance of our anomaly detection method, we use two evaluation metrics which are Mean Absolute Error (MAE) and Intersection over Union (IoU). MAE quantifies the average difference between the original signal (X_i) and reconstruction signal $(\hat{X}_i)$ over N samples where lower MAE value indicates more accurate signal reconstruction by the autoencoder.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} | X_i - \hat{X}_i | \tag{5}$$

IoU analyze the overlap area between the detected anomaly and the ground truth anomaly. A is the set of detected anomaly indices and G is the set of ground truth anomaly indices. Higher IoU value indicates better localization and stronger agreement between the detected anomalies and the ground truth.

$$IoU = \frac{A \cap G}{A \cup G} \tag{6}$$

In general, MAE is a widely used metric for evaluating reconstruction quality in regression and signal processing tasks, while IoU is commonly used to evaluate the accuracy of predicted regions or segments by comparing the predicted area and the ground truth area in applications such as image segmentation or event detection.

IV. RESULTS AND DISCUSSION

Anomaly detection evaluation was conducted by applying a separate autoencoder model to each plant species. In the initial phase of the experiment, we compared preprocessing with and without detrending to determine the most suitable approach for our dataset.

Based on the results presented in Table 2, the comparison shows that applying detrending during preprocessing generally leads to clear improvements in model performance. Therefore, detrending was selected as the standard preprocessing method for the remainder of this study. After establishing the effective preprocessing strategy, we proceeded to evaluate the overall performance of the proposed model across the complete dataset.

TABLE II. DETREND COMPARISON

Plant Species	**Without Detrend (IoU)**	**With Detrend (IoU)**
Ocimum	0.558	0.615
Plectranthus	0.669	0.701
Solanum	0.530	0.670

The results in Table 3 show that, under identical preprocessing pipelines and training procedures, reliable abnormality detection was achieved only for the species *Ocimum, Plectranthus, Ruta, and Solanum* species. These four species attained notably higher average IoU and IoU > 0.5 scores, indicating better alignment between the detected and ground truth anomaly regions (Fig. 5). Meanwhile, species with lower IoU scores often also show lower MAE values, which may suggest either a lack of significant signal changes in response to stimuli or the presence of high signal noise (Fig. 6). Both of which make it difficult for the model to distinguish between normal and anomalous patterns.

TABLE III. EACH SPECIES PERFORMANCE EVALUATION

Plant Species	**MAE Loss**	**Average IoU**	**IoU > 0.5**
Araucaria	0.0476	0.435	57.1%
Cyperus	0.0185	0.426	25.0%
Hedera	0.0511	0.050	0%
Mentha	0.0681	0.245	7.7%
Ocimum	0.0635	**0.615**	**68.4%**
Plectranthus	0.0662	**0.701**	**94.1%**
Ruta	0.0374	**0.631**	**84.6%**
Salvia	0.0500	0.348	33.3%
Solanum	0.0563	**0.670**	**84.6%**

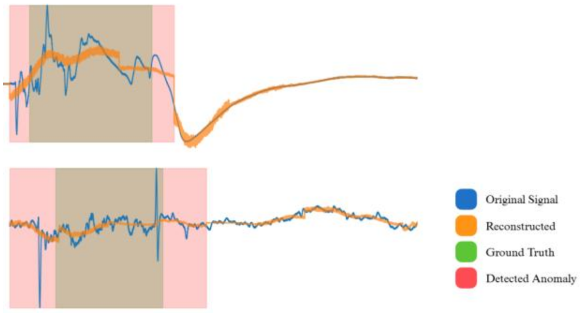

Fig. 5. Success Inference Result.

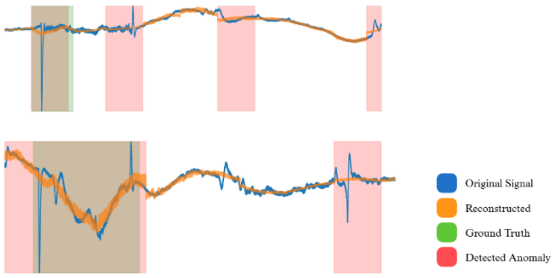

Fig. 6. Failed Inference Result.

It was also observed that the dataset contained uncertain data, with stimulus-like signals appearing outside the labeled stimulus regions and non-stimulus-like signals present within the stimulus regions. (Fig. 6). This uncertainty is observed across all species. However, quantitative analysis and our observations reveal that *Ocimum*, *Plectranthus*, *Ruta*, and *Solanum* exhibit a lower incidence of uncertain data compared to other species. This inconsistency can cause the model to learn false peaks and incorrectly classify them as normal. As a result, during prediction, the model may fail to detect actual anomalies or misclassify normal-like patterns within the stimulus area, leading to lower IoU scores. While the autoencoder was effective in identifying anomaly patterns in the signals, its performance was highly dependent on the species-specific signal responses to both stimulus and normal conditions.

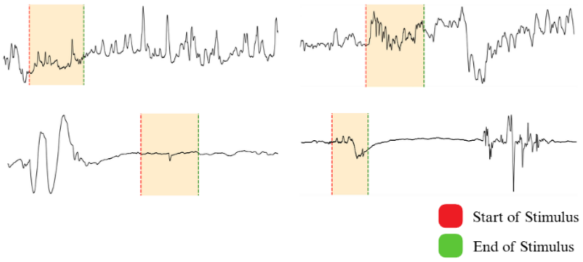

Fig. 7. Uncertain Data.

Overall, the findings suggest that the success of autoencoder-based anomaly detection can vary considerably across plant species. Factors such as signal quality, inherent physiological response, and the amount of available training data likely contribute to these differences.

V. CONCLUSION

This study highlight the effectiveness of an autoencoder-based approach for detecting abnormal pattern in plant electrophysiological signal triggered by external stimuli. By using preprocessing methods such as detrending and normalization, the model is able to detect anomaly patterns in the signal for *Ocimum, Plectranthus, Ruta, and Solanum* plant species. As for the 4 plant species mentioned, the model achieves average IoU scores from 0.615 to 0.701 and IoU $>$ 0.5 scores from 68.4% to 94.1% indicating a successful alignment between the ground truth and predicted anomaly area. For the other species, the model exhibited much lower average IoU and IoU $>$ 0.5 scores ranging from 0.050 to 0.435 and 0% to 57.1%. This may indicate a weaker and uncertain responses or challenge in signal quality and annotation accuracy. These result shows that while autoencoder approach can detect abnormalities in some plant species, their performance is highly correlated on species-specific signal pattern. Future work should aim to improve detection in underperforming species by incorporating additional signal features and enhancing the quality of data annotation and collection strategies.

VI. ACKNOWLEDGMENT

This work was supported by funding from the National Science and Technology Council (NSTC), Taiwan, under grants NSTC 114-2923-E-008-003-MY3

REFERENCES

[1] V. A. Vodeneev, L. A. Katicheva, and V. S. Sukhov, “Electrical signals in higher plants: Mechanisms of generation and propagation,” *Biophysics (Russian Federation)*, vol. 61, no. 3, pp. 505–512, May 2016, doi: 10.1134/S0006350916030209.

[2] T. T. H. Phan, T. M. V. Ngo, and H. P. Phan, “Flexible Mechanical Sensors for Plant Growth Monitoring: An Emerging Area for Smart Agriculture,” *Sensors 2024, Vol. 24, Page 7995*, vol. 24, no. 24, p. 7995, Dec. 2024, doi: 10.3390/S24247995.

[3] R. Chalapathy and S. Chawla, “Deep Learning for Anomaly Detection: A Survey,” Jan. 2019, Accessed: May 23, 2025. [Online]. Available: https://arxiv.org/pdf/1901.03407

[4] D. Gong *et al.*, “Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection,” *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1705–1714, Oct. 2019, doi: 10.1109/ICCV.2019.00179.

[5] M. Jamjoom, A. Elhadad, H. Abulkasim, and S. Abbas, “Plant Leaf Diseases Classification Using Improved K-Means Clustering and SVM Algorithm for Segmentation,” *Computers, Materials & Continua*, vol. 76, no. 1, pp. 367–382, Jun. 2023, doi: 10.32604/CMC.2023.037310.

[6] S. M. Javidan, A. Banakar, K. A. Vakilian, and Y. Ampatzidis, “Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning,” *Smart Agricultural Technology*, vol. 3, Feb. 2023, doi: 10.1016/J.ATECH.2022.100081.

[7] S. Schmidl, F. Naumann, and T. Papenbrock, “AutoTSAD: Unsupervised Holistic Anomaly Detection for Time Series Data,” *Proceedings of the VLDB Endowment*, vol. 17, no. 11, pp. 2987–3002, Jul. 2024, doi: 10.14778/3681954.3681978.

[8] P. Liznerski, L. Ruff, R. A. Vandermeulen, B. J. Franks, M. Kloft, and K. R. Müller, “Explainable Deep One-Class Classification,” *ICLR 2021 - 9th International Conference on Learning Representations*, Jul. 2020, Accessed: May 23, 2025. [Online]. Available: https://arxiv.org/pdf/2007.01760

[9] D. Madariaga *et al.*, “A library of electrophysiological responses in plants - a model of transversal education and open science,” *Plant Signal Behav*, vol. 19, no. 1, Dec. 2024, doi: 10.1080/15592324.2024.2310977.

[10] T. Aust, E. Buss, F. Mohr, and H. Hamann, "Automated Phytosensing: Ozone Exposure Classification Based on Plant Electrical Signals," *2025 IEEE Symposia on Computational Intelligence for Energy, Transport and Environmental Sustainability, CIETES 2025*, 2025, doi: 10.1109/CIETES63869.2025.10995087.

[11] H. Li, N. Fotouhi, F. Liu, H. Ji, and Q. Wu, "Early detection of dark-affected plant mechanical responses using enhanced electrical signals," *Plant Methods*, vol. 20, no. 1, pp. 1–14, Dec. 2024, doi: 10.1186/S13007-024-01169-4/TABLES/7.

[12] M. Mudrilov, M. Ladeynova, Y. Vetrova, and V. Vodeneev, "Analysis of the Mechanisms Underlying the Specificity of the Variation Potential Induced by Different Stimuli," *Plants*, vol. 13, no. 20, p. 2896, Oct. 2024, doi: 10.3390/PLANTS13202896/S1.

[13] H. Wen *et al.*, "Mutually reinforcing and transpiration-dependent propagation of H2O2 and variation potential in plants revealed by fiber organic electrochemical transistors," *Innovation*, May 2025, doi: 10.1016/J.XINN.2025.100800.

[14] E. Ghasemi, E. Ebrahimie, and A. Niazi, "Machine learning for early detection of plant viruses: Analyzing post-infection electrical signal patterns," *Smart Agricultural Technology*, vol. 9, Dec. 2024, doi: 10.1016/J.ATECH.2024.100668.

[15] A. C. U. Furch *et al.*, "Transformation of flg22 perception into electrical signals decoded in vasculature leads to sieve tube blockage and pathogen resistance," *Sci Adv*, vol. 11, no. 9, p. eads6417, Feb. 2025, doi: 10.1126/SCIADV.ADS6417.

[16] R. Kopel *et al.*, "No time for drifting: Comparing performance and applicability of signal detrending algorithms for real-time fMRI," *Neuroimage*, vol. 191, pp. 421–429, May 2019, doi: 10.1016/J.NEUROIMAGE.2019.02.058.

[17] A. de Cheveigné and D. Arzounian, "Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data," *Neuroimage*, vol. 172, p. 903, May 2018, doi: 10.1016/J.NEUROIMAGE.2018.01.035.

[18] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," *ACM International Conference Proceeding Series*, vol. 02-December-2014, pp. 4–11, Dec. 2014, doi: 10.1145/2689746.2689747.