Electrophysiological Signal Study: Measuring, Analyzing, and Classifying Plant Responses

Bhoumik Patidar (22110049), Birudugadda Srivibhav (22110050)
Computer Science And Engineering, Indian Institute Of Technology Gandhinagar

---

Project Supervisor: Professor Subramanian Sankaranarayanan, Department Of Biological Sciences And Engineering
The report was submitted on 21-11-2023

This project delves into the interdisciplinary realm of biology, electrical engineering, and computer science to explore and understand the electrophysiological responses of tobacco plants to stimuli, particularly wound-induced variation potentials. The endeavour involved the cultivation of tobacco plants, the development of an electrophysiological measurement device, and the use of a DAQ system for data collection. Despite challenges faced in the physical realization of the device, experiments were conducted to measure variation potentials resulting from plant wounding. Subsequently, collected data was processed and analysed using signal processing techniques and a Support Vector Machine (SVM) model was developed to identify regions within the signals corresponding to plant wound responses. This project serves as an interdisciplinary approach, demonstrating the integration of experimental biology, electrical engineering, and computational methods to study and interpret plant electrophysiology.

Experimental Plant Distribution And Growing Conditions

Nicotiana tabacum plants were cultivated in a controlled laboratory environment to ensure consistent and diverse data collection. The experiment involved 32 plants, segregated into two batches of 16 plants each. These batches were grown separately, introducing a two-month interval between their initiation to encompass a varied age range within the dataset.

1. **Growing Medium Preparation:** To facilitate healthy plant growth, a specific soil mixture was meticulously prepared. This mixture comprised perlite, soilrite, vermiculite, and soil in a precise ratio of 1:1:3:1, offering an optimal environment for Nicotiana tabacum development.
2. **Sterilization Process and Maintenance of Sterile Environment:** The soil underwent rigorous sterilization by exposure to a temperature of 120 degrees Celsius, effectively eliminating all potential microorganisms, pathogens, and pests that could hinder plant growth. Additionally, the pots utilized for planting were meticulously cleaned with ethanol, ensuring a controlled and sterile environment conducive to the optimal growth of Nicotiana tabacum.

The plants grew in a chamber that provided controlled conditions for temperature, humidity, and light, optimizing plant development and ensuring scientific standards.

Signal Measurement Methodology

1. **Electrode Preparation:** To measure plant signals using invasive methodology, Ag/AgCl electrodes were prepared. A 4-cm-long silver wire was cut and chloridized on one end in a 0.1 M HCl solution to create a layer of AgCl, crucial for signal detection. The chloridation process involved connecting the electrode to the anode of a 1.5-V battery and immersing its end in the solution for a short duration, resulting in the formation of a visible, dark coating of AgCl.

2. **Experimental Setup and Signal Measurement:** In the experimental setup, one electrode was inserted near the leaf's base, and the other was grounded. When the leaf was wounded, electrical recordings were captured using the Ag/AgCl electrodes. This invasive methodology facilitated accurate measurement of plant electrical signals, particularly the variation potential induced by leaf injury.

Attempt At Signal Processing Circuit Design

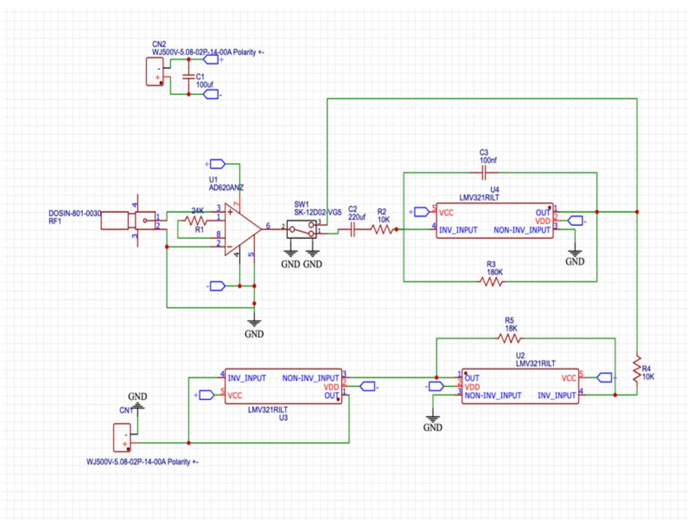
Circuit Design and Prototyping: The circuit design for filtering and amplifying received signals from the probes was planned and executed in two stages—initial design in LTSpice and subsequent PCB prototyping using EasyEDA.

Circuit Specifications: The designed circuit aimed for a total amplification of 72 times and signal filtering within the 0.01 Hz to 10 Hz range. The structure comprised:

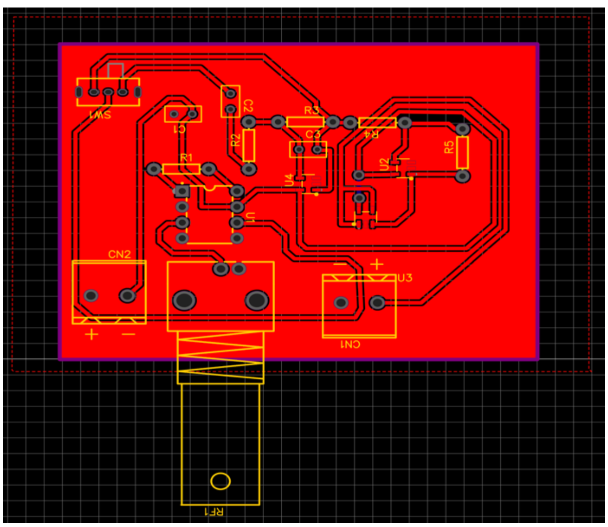
1. **AD620 Amplifier (2x Amplification):** The initial phase utilized an AD620 amplifier for a 2x amplification.
2. **Frequency Filtering and 18x Amplification:** Following the AD620, the signal underwent frequency filtering within the specified range and experienced an additional 18x amplification.
3. **2x Amplification Phase:** An intermediate 2x amplification stage.
4. **Buffer Phase:** Finally, the signal passes through a buffer phase before entering the Analog-to-Digital Converter (ADC) for computer data collection.

Purpose of Each Circuit Phase:

1. **AD620 Amplifier (2x Amplification):** To provide an initial amplification to the received weak signals for better processing.
2. **Frequency Filtering and 18x Amplification:** To filter out unwanted frequencies and further amplify the filtered signal.
3. **2x Amplification Phase:** An additional amplification stage to enhance signal strength.
4. **Buffer Phase:** To ensure proper impedance matching and prevent signal loss before feeding into the ADC for digitization.



Circuit Design

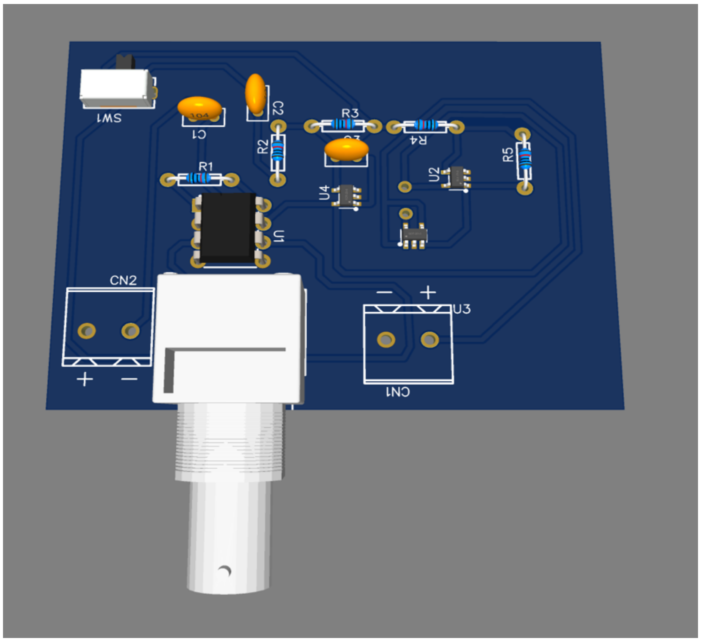


PCB Layout

Outcome and Analysis: Upon actualizing the circuit, it encountered significant noise issues, rendering it non-functional. While the circuit design performed well in the LTSpice simulation, its physical implementation failed due to excessive noise interference.

Possible Reasons for Failure:

1. **Component Tolerances:** Variations in component characteristics from their nominal values in the physical circuit.
2. **Interference and Crosstalk:** Noise introduced from external sources or inadequate isolation between circuit components.
3. **Grounding Issues:** Improper grounding leads to noise coupling into the circuit.
4. **PCB Layout and Trace Design:** Suboptimal layout or traces causing interference or signal degradation.



A 3D visualization of the circuit

Data Acquisition Methodology

Electrode Connection and Signal Input: The electrodes were soldered to wires, establishing a reliable connection for analogue signal transmission. These wired electrodes were then directly interfaced as analogue inputs to the NI DAQ 6009 device, serving as the primary conduit for signal transmission from the electrodes to the data acquisition hardware.

Utilization of NIDAQMX Library for Signal Handling: To manage and interpret the incoming analogue signals, the NIDAQMX library in Python was utilized. This powerful library facilitated the acquisition, processing, and interpretation of the analogue signals received from the electrodes connected to the DAQ 6009. It provided a streamlined approach for handling and storing the acquired electrophysiological data on the computer.

Sampling Rate: The data acquisition process operated at a sampling rate of 100 samples per second. This rate ensured the capture of signal dynamics at a high temporal resolution, enabling detailed analysis and interpretation of the electrophysiological responses.

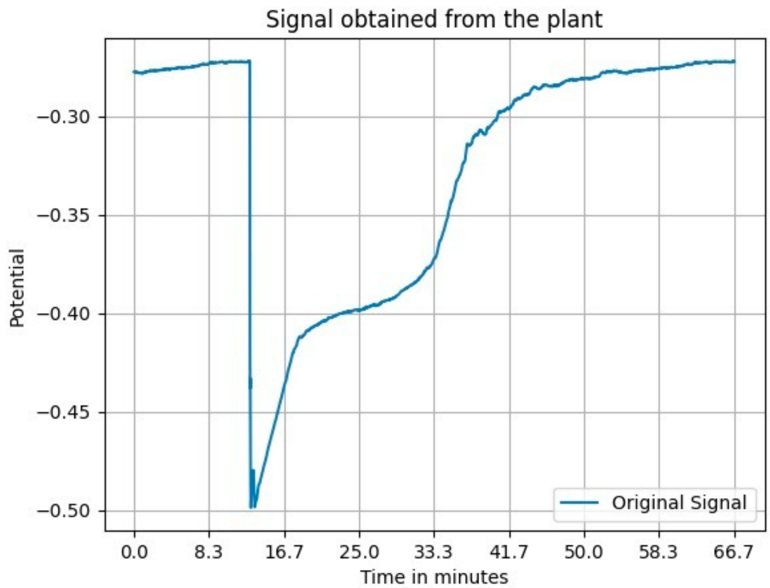
Data Collection and Pre-processing

Variation Potential Signal Description: The variation potential signal exhibits a weak amplitude, reaching a peak of a few tens of millivolts. Upon plant injury, the leaf's surface potential experiences a sudden drop over a few seconds, gradually returning to a constant normal value within 15 to 45 minutes. In a healthy state, the signal remains relatively constant, yet this value varies among different plants and even for the same plant on different days. This extracellular measurement captures potential changes at the intracellular level indirectly.

Data Description: Over 50 hours of plant electrical signal data were collected, including 36 variation potential measurements, an equal number of baseline data representing a healthy plant state, and additional noisy data, likely due to external factors like wind or shaking, which shouldn't be classified as variation potential by the ML model.

Digital Data Filtering: A frequency filter was utilized to eliminate frequencies above a certain value. The subsequent filtering process involved a combination of steps executed in code:

1. **Butterworth Filter:** Applied a low-pass Butterworth filter to the raw voltage data.
2. **Median Filtering:** Employed median filtering on the filtered voltage signals.
3. **Savitzky-Golay Filter:** Utilized the Savitzky-Golay filter to further smooth the signals. These steps aimed to enhance the signal quality and reduce noise, resulting in smoother, more manageable data for model input.



Exploratory Data Analysis

Visualization:

Purpose: Visualized variation potential signals alongside normal baseline and noisy signals.

Approach: Graphical representation to discern patterns and differences between variation potential and other signal types.

Insights: Identified distinctive patterns characterizing variation potential signals against normal baseline and noisy signals.

Statistical Analysis:

Objective: Quantified differences between variation potential and other signal types.

Tasks: Applied statistical measures to differentiate statistical properties of variation potential from normal baseline and noisy signals.

Outcomes: Enabled quantification of distinctions among signal types for subsequent classification modelling.

Data Cleaning Confirmation:

Objective: Ensured data consistency for accurate classification modelling.

Steps Undertaken: Checked for data consistency post-pre-processing to maintain reliability.

Impact: Verified data integrity, crucial for delineating variation potential from normal baseline and noise during classification.

The Algorithm

Introduction: The algorithm is designed to analyse plant signal data, specifically targeting the detection of distinct phases in the signal corresponding to plant injury events. The plant signal exhibits two crucial phases following an injury: a sudden drop in potential, signifying the wounding time, followed by a gradual recovery to the baseline potential. The primary goal of this algorithm is twofold: first, to pinpoint the time of injury onset, and second, to delineate the duration of the variation potential**.**

Overall Algorithm Overview: This algorithm comprises two distinct models, each addressing a specific phase in the plant signal:

Model 1: Detection of the First Phase (Wounding Time)

Data Preparation:

Training Segments: Segments of signal data, each comprising 500 samples (~8 seconds), containing:

a. Wounding segments (35 training examples) showing a sudden drop.
b. Healthy plant signal segments (68 training examples) without sudden drops.

Feature Extraction:

Features: Extracted features from each 500-sample segment:

a. Max-min difference
b. Last value minus initial value
c. Variance
d. Interquartile Range (IQR)

Feature Scaling:

Standardization: Applied standard scaling using the StandardScaler to maintain consistent feature scaling.

Labelling and Training:

Target Labels (y_train): Assigned labels where 1 represents sudden drop segments, and 0 represents non-sudden drop segments.

Feature-Label Pairing (X_train): Created dataset pairing extracted features (X_train) with target labels (y_train).

Model Training:

Support Vector Machine (SVM): Trained an SVM model using the RBF kernel on the dataset (X_train, y_train) to learn the patterns distinguishing sudden drop segments from non-sudden drop segments.

This model aims to identify the first phase of the signal corresponding to the sudden drop or wounding time by training an SVM classifier on extracted features from segments of plant signal data.

Model Validation And Evaluation

1. **High Accuracy and Recall:**
- Achieved a high accuracy of approximately 96.77%.
- Demonstrated perfect recall, indicating it effectively identified all wounded segments.
2. **Good Precision and F1-Score:**
- Showed commendable precision of about 92.31%.
- Presented a strong F1-score of approximately 96.0%, signifying a balance between precision and recall.
3. **Cross-Validation Performance:**
- Displayed consistent accuracy across different folds, averaging 94.48%.

Model 2: Detection of recovery phase

Data Preparation:

Segments: Training on segments of length 5000 elements (~80 seconds) comprising:

1. Recovery segments (6681 samples) represent the recovery phase.
2. Not recovery segments (5410 samples) denoting other phases.

Feature Extraction:

Features: Extracted features from each 5000-element segment:

1. Time gap from the wounding time in the input signal (predicted using Model 1).
2. Statistical measures:
3. Max, min, max-min difference.
4. Last value, initial value, last value-initial value.
5. Variance and Interquartile Range (IQR).

Feature Scaling:

Standardization: Applied standard scaling using the StandardScaler to maintain consistent feature scaling.

Labelling and Training:

Target Labels (y_train): Assigned labels where 1 represents the recovery phase, and 0 represents non-recovery segments.

Feature-Label Pairing (X_train): Created dataset pairing extracted features (X_train) with target labels (y_train).

Model Training:

Support Vector Machine (SVM): Trained an SVM model using the RBF kernel on the dataset (X_train, y_train)

Model Validation And Evaluation:

1. **High Accuracy and Recall:**
 - Attained an accuracy of around 96.39%.
 - Showcased a robust recall score of approximately 97.35%, suggesting effective identification of the variation phases.
2. **Balanced Precision and F1-Score:**
 - Demonstrated a good precision of about 96.26%.
 - Maintained a strong F1-score of around 96.80%, indicating a balance between precision and recall.
3. **Cross-Validation Performance:**
 - Exhibited consistent accuracy across folds, averaging 96.44%.

Algorithm Workflow: Detection of Injury Onset and Variation Potential Duration

The algorithm is designed with a two-fold objective: identifying the onset of injury and determining the duration of the variation potential within the input signal data.

Part 1: Detection and Confirmation of Wounding Time

- **Model 1 Analysis:**

  Analyses 500-sample segments using Model 1 to detect wounded signal segments.

- **Wounding Confirmation Conditions:**

  Condition 1: Identifies the detected segment as wounded; ensures that the subsequent 20 segments are not classified as wounded.

  Condition 2: Validate the subsequent five segments as recovery phase segments using Model 2.

- **Wounding Time Determination:**

  If both conditions are met, the algorithm confirms the detected segment as wounded, and the timestamp for that segment is recorded as the time instant for wounding.

Part 2: Detection of Recovery Phase Duration

- **Model 2 Analysis:**

  Analyses 5000-length segments using Model 2 to discern recovery phase segments.

- **Recovery Phase Duration Determination:**
  i. Utilizes the previously identified wounding time as the start of the recovery phase.
  ii. Sequentially evaluates 5000-length segments using Model 2, starting from the wounding time, to identify the duration of the recovery phase.
  iii. Marks the end of the recovery phase when five consecutive 5000-length segments are classified as 'not recovery phase' by Model 2.
- **Fallback Procedure:**

  If the initial conditions for confirming wounding are not met, the algorithm iterates through subsequent 500-length segments, seeking a wounded segment that fulfils the conditions required for detecting the start of the recovery phase.

This algorithm effectively leverages Model 1 to pinpoint the time of injury onset and employs Model 2 to delineate the duration of the variation potential within the input signal data, ensuring the fulfilment of specific conditions for accurate detection and confirmation.

	Model - 1 Performance	Model - 2 Performance
Accuracy	96.77%	96.39%
Recall	100%	97.35%
Precision	92.31%	96.26%
F1-score	96.0%	96.80%
Cross-Validation	94.48%	96.44%

End-to-End Testing on Unseen Data:

- Overall Algorithm Performance:

The algorithm displayed strong performance on unseen signal data, accurately determining both the onset of injury and the recovery time range.

Task 1 Success Rate:

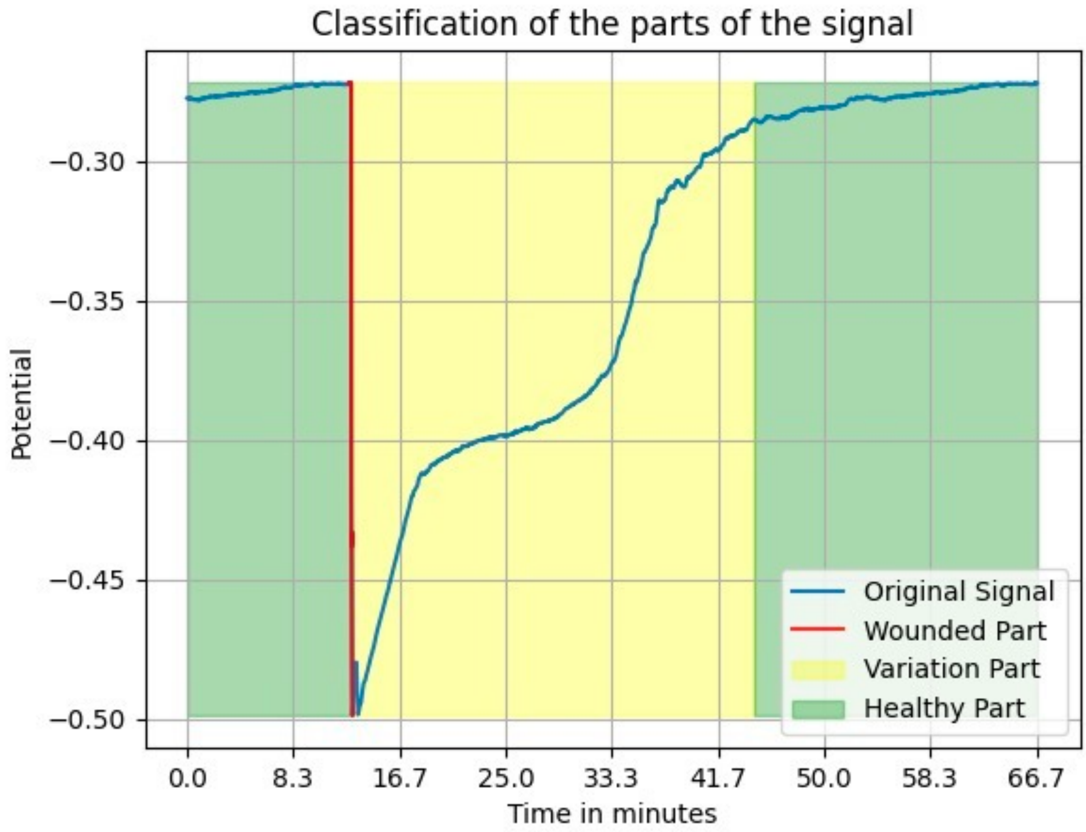
Achieved a perfect success rate of 100% in accurately identifying the onset of injury.

Task 2 Success Rate:

Predicted the variation potential duration with good precision, indicating effective delineation of the recovery phase.

Overall Success Metric:

The combined success of both Task 1 and Task 2 showcases the algorithm's robustness in accurately detecting injury onset and delineating variation potential durations on unseen data.



Acknowledgements

We extend our gratitude to:

- Prof. Subramanian Sankaranarayanan for allowing us to do this project under his supervision. This project was of great learning for us. We are extremely thankful to him for his continuous support and encouragement throughout the project.
- Prof. Archana Kumari, GBU, for her invaluable guidance in plant electrophysiology.
- Prof. Rajendran, for his expertise in circuit design.
- PhD Scholar Hemal Bhalla for his support and ongoing assistance with plant cultivation.
- Postdoc Charlie Kaushal for her dedicated support in plant cultivation.
- The staff of the Electrical Engineering and Biological Sciences and Engineering labs.