SAUVOLA SEGMENTATION AND SUPPORT VECTOR MACHINE-SALP SWARM ALGORITHM APPROACH FOR IDENTIFYING NUTRIENT DEFICIENCIES IN CITRUS RETICULATA LEAVES

LIA KAMELIA

ASIA e UNIVERSITY
2024

SAUVOLA SEGMENTATION AND SUPPORT VECTOR MACHINE-SALP SWARM ALGORITHM APPROACH FOR IDENTIFYING NUTRIENT DEFICIENCIES IN CITRUS RETICULATA LEAVES

LIA KAMELIA

A Thesis Submitted to Asia e University in
Fulfilment of the Requirements for the
Degree of Doctor of Philosophy

March 2024

ABSTRACT

Machine learning and image processing methods can effectively detect nutrient deficiencies in citrus trees, addressing the challenge of accurately identifying shortages that can impair crop health and productivity. Traditional methods often rely on expert visual assessments, which are labour-intensive, subjective, and time-consuming. The proposed method integrates colour and texture feature-based image analysis with machine learning algorithms for classification. The process begins with acquiring image data, which is categorized into four classes: nitrogen (N) deficiency, phosphorus (P) deficiency, potassium (K) deficiency, and normal. In total, 1,200 images are collected. Next, file sizes are reduced using lossless compression methods, achieving a 96.99% reduction. The second phase involves image segmentation using the Sauvola method. Following this, colour and texture feature extraction is performed. Colour features are extracted in the Hue (H), Saturation (S), and Value (V) colour space, while texture features are obtained using the Grey-Level Co-Occurrence Matrix (GLCM) method. This combination of colour and texture features results in various metrics, including mean, dissimilarity, skewness, angular second moment, variance, entropy, maximum probability, contrast, correlation, energy, and homogeneity, which are used for classification. Both Support Vector Machine (SVM) and Artificial Neural Network (ANN) methods are compared for classification. The Sauvola method combined with ANN achieves the highest accuracy of 93.75%. In the next phase, the datasets are optimized using the Salp Swarm Algorithm (SSA), which improves classification accuracy. With SSA optimization, the Sauvola method combined with SVM reaches an accuracy of 99.58%, surpassing other methods that use image processing and ANN classification. Expert validation is utilized to evaluate and validate the effectiveness of the proposed method and confirm the system's accuracy at 95%. Integrating SSA and SVM machine learning algorithms improves decision-making processes, leading to better crop yield through early detection and timely nutrient management. It ensures that plants receive the necessary nutrients for optimal growth and development.

Keywords: Citrus leaves, classification, Sauvola segmentation, optimization, Salp Swarm Algorithm

APPROVAL

This is to certify that this thesis conforms to acceptable standards of scholarly presentation and is fully adequate, in quality and scope, for the fulfilment of the requirements for the degree of Doctor of Philosophy

The student has been supervised by: **Prof Ts Dr Titik Khawa Abdul Rahman, Dean SST & SSF & Senior Advisor, QARA**

The thesis has been examined and endorsed by:

Prof. Dr. Ku Ruhana Bt Ku Mahamud,
Universiti Utara Malaysia (UUM)
Examiner 1

Profesor Dr Nooritawati Md Tahir,
Uiniversiti Teknologi Mara (UiTM)
Examiner 2

This thesis was submitted to Asia e University and is accepted as fulfilment of the requirements for the degree of Doctor of Philosophy.

Professor Dr. Siow Heng Loke
Asia e University
Chairman, Examination Committee
(8 March 2024)

DECLARATION

I hereby declare that the thesis submitted in fulfilment of the PhD degree is my own work and that all contributions from any other persons or sources are properly and duly cited. I further declare that the material has not been submitted either in whole or in part, for a degree at this or any other university. In making this declaration, I understand and acknowledge any breaches in this declaration constitute academic misconduct, which may result in my expulsion from the programme and/or exclusion from the award of the degree.

Name: Lia Kamelia

Signature of Candidate: **Date**: 8 March 2024

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious, the Most Merciful. All praise and gratitude are due to Allah SWT for His infinite blessings and guidance that have enabled me to complete this thesis.

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Prof. Titik Khawa Abdul Rahman, at Asia e University, for her invaluable support, insightful guidance, and constant encouragement throughout the entire process of my research and thesis writing. Her expertise and dedication have been instrumental in completing this work. I want to extend a special thanks to the Dean of the Faculty of Science and Technology and all the lecturers in the Department of Electrical Engineering at UIN Sunan Gunung Djati Bandung for their valuable input and assistance. I am deeply thankful to my colleagues and UIN Sunan Gunung Djati students for their continuous support and encouragement.

I owe my deepest gratitude to my beloved parents, who have passed away during my study. Their love, prayers, and sacrifices have been my greatest source of inspiration. May Allah grant them Jannah. My heartfelt thanks go to my dear husband, Cece Apandi, whose unwavering love, patience, and support have been my pillars of strength. To my wonderful children, Haziq Sulthan Alifandi and Kalila Sayyida Alifandi, thank you for your love and understanding, which are my source of joy and motivation.

Finally, I would like to extend my gratitude to all those who have supported me in any respect during the completion of this thesis. May Allah reward all your help with better rewards.

TABLE OF CONTENTS

ABSTRACT		**ii**
APPROVAL		**iii**
DECLARATION		**iv**
ACKNOWLEDGEMENTS		**vi**
TABLE OF CONTENTS		**vii**
LIST OF TABLES		**x**
LIST OF FIGURES		**ii**
LIST OF ABBREVIATION		**vii**
CHAPTER 1	**INTRODUCTION**	**1**
1.0	Background of the Study	1
1.1	Problem Statement	8
1.2	Research Objectives	9
1.3	Research Questions	9
1.4	Research Hypotheses	15
1.5	Justification and Significance of the Study	15
1.6	Theoretical Contributions	17
1.7	Practical Contributions	18
1.8	Contribution to Methodology	19
1.9	Scope Work and Limitation	19
1.10	Chapter Summary	20
CHAPTER 2	**REVIEW OF LITERATURE**	**22**
2.0	Introduction	22
2.1	Nutrient Deficiency in Citrus	22
	2.1.1 Nitrogen Deficiency	25
	2.1.2 Phosphorus Deficiency	26
	2.1.3 Potassium Deficiency	27
2.2	The Methods to Detect Nutrient Deficiency for Plant	29
2.3	Image Processing	32
	2.3.1 Image Pre-Processing Technique	33
	2.3.2 Image Segmentation	40
	2.3.3 Feature Extraction	46
2.4	Classification Technique	52
	2.4.1 Artificial Neural Network	54
	2.4.2 Support Vector Machine (SVM)	56
2.5	Optimisation of ANN and SVM	61
2.6	Application of AI in Agricultural Industries	70
2.7	Validation Using Expert’s Opinion	76
2.8	Chapter Summary	78
CHAPTER 3	**METHODOLOGY**	**80**
3.0	Introduction	80
3.1	Research Design	80
3.2	Research Framework	83

3.3 Development of Image Processing Technique for Feature Extraction in Order to Detect Nutrient Deficiency in Citrus Leaves 89
3.3.1 Initial Sample Treatment Before Image Acquisition 90
3.3.2 Citrus Leaves Image Pre-Processing 95
3.3.3 Image Segmentation Using Canny and Sauvola Methods 100
3.3.4 Feature Extraction of Image based on the Colour and Texture 106
3.4 The Development of a Classification System from Extracted Image 113
3.4.1 The Classification System Using SVM Algorithm based on HSV-GLCM Parametric 114
3.4.2 The Classification System Using ANN Algorithm based on HSV-GLCM Parametric 116
3.5 The Development of the Optimisation Algorithm Using the Salp Swam Algorithm 123
3.5.1 The Optimisation of ANN Using SSA 126
3.5.2 The Optimisation of SVM Using SSA 130
3.6 Analysis of Expert Validation Using Judge Opinion for Implemented System 135
3.7 Comparative Study 140
3.8 Chapter Summary 143

CHAPTER 4 RESULTS AND DISCUSSION 145

4.0 Introduction 145
4.1 Experimental Design 145
4.2 Texture and Colour Feature Extraction to Detect the Nutrient Deficiency in Citrus Leaves 148
4.2.1 Results of Initial Treatment 148
4.2.2 Results of Image Acquisition as RGB Image File 150
4.2.3 The Size Image Reduction Using Lossless Compression 151
4.2.4 The Process of Image Enhancement with Contrast Stretching Methods 153
4.2.5 The Image Segmentation Using Canny and Sauvola's Methods 160
4.2.6 Comparison Results in the Canny and Sauvola Segmentation Process 164
4.2.7 Texture and Colour Extraction to Detect the Nutrient Deficiency in Citrus Leaves 167
4.2.8 Analysis and Visualization of Colour and Texture Features 170
4.3 The Classification of Nutrient Deficiency Based on Colour and Texture of Citrus Leaves 175
4.3.1 Data Preparation for the Classification 176
4.3.2 The Development of Artificial Neural Network for the Classification of Nutrient Deficiency in Citrus Leaves 177
4.3.3 The Development of a Support Vector Machine for the Classification of Nutrient Deficiency in Citrus Leaves 198
4.4 Optimisation of ANN and SVM Classification Using the Salp Swarm Algorithm 210
4.4.1 Optimisation of ANN Using SSA 211

4.4.2 The Optimisation of SVM Using SSA 230
4.4.3 Comparison of ANN and SVM Models Before and After Optimisation with SSA 244
4.5 The Expert Validation of the Classification System 252
4.6 The Comparative Study with Benchmark Dataset 260
4.7 Chapter Summary 263

CHAPTER 5 CONCLUSION, LIMITATION AND RECOMMENDATIONS 265

5.0 Conclusion 265
5.1 Limitations of Research 267
5.2 Recommendations for Future Research 267
REFERENCES 269
APPENDICES 287
Appendix A- Data Set for the Expert Validation 287

LIST OF TABLES

Table		Page
1.1	Alignment between research objective, problem statement and research questions	11
2.1	Previous research and gap	73
3.1	Research framework	84
3.2	Type of crop treatment	91
3.3	The composition of nutrient solution for various conditions	92
3.4	Formula of accuracy, precision, recall and F1 score	122
4.1	Comparison of macronutrient deficiencies syndrome	149
4.2	PSNR mean for each category	155
4.3	The image's SSIM after compression and contrast stretching	157
4.4	The average calculation of HSV-GLCM parameters from Canny data	168
4.5	Colour-texture extraction selected output	169
4.6	The output of the data split process	176
4.7	The detailed number of data	177
4.8	Output of the 5-fold validation for ANN - Canny segmented data	178
4.9	Classification results using ANN from Canny segmented data	184
4.10	ROC AUC for ANN using Canny segmented data	184
4.11	Output of the 5-fold validation for ANN-Sauvola segmented data	189
4.12	Classification results using ANN from Sauvola segmented data	194
4.13	ROC AUC for ANN using Sauvola segmented data	194
4.14	Output of the 5-fold validation for SVM - Canny segmented data	198
4.15	Classification results using SVM from Canny segmented data	199
4.16	ROC AUC for SVM using Canny segmented data	200

4.17	Output of the 5-fold validation for SVM – Sauvola segmented data	204
4.18	Classification results using SVM from Sauvola segmented data	206
4.19	ROC AUC for SVM using Sauvola segmented data	206
4.20	Best parametric for ANN using SSA using Canny segmented data	211
4.21	Classification results using SSA-ANN from Canny segmented data	216
4.22	ROC AUC for SSA-ANN using Canny segmented data	218
4.23	Best parametric for ANN using SSA using Sauvola segmented data	222
4.24	Classification results using SSA- ANN from Sauvola segmented data	226
4.25	ROC AUC for SSA-ANN using Sauvola segmented data	227
4.26	Best hyperparameters for SSA-SVM from Canny segmented data	230
4.27	Classification results using SSA-SVM from Canny segmented data	233
4.28	ROC AUC for SSA-SVM using Canny segmented data	236
4.29	Best hyperparameters for SSA-SVM from Sauvola’s segmented data	239
4.30	Classification results using SSA-SVM from Sauvola segmented data	241
4.31	ROC AUC for SSA-SVM using Sauvola segmented data	242
4.32	Recapitulation data of all processes	246
4.33	Output results of selected unseen data	253
4.34	Results of classification for unseen data using the best model	255
4.35	Expert's judgment versus system classification	257
4.36	Number of digital images by nutritional deficiency of coffee leaves from the Coleaf-DB dataset	260
4.37	Confusion matrix for Co-leaf dataset using the proposed model	262

LIST OF FIGURES

Figure		Page
1.1	Citrus reticulata Sp	2
2.1	Visual differences in nutritional deficiency	24
2.2	The different types of leaves at NPK deficiencies	25
2.3	The general formation of a classification system	54
2.4	(a) Straight lines distinct two classes. (b) The best line. The area between the classes is a margin	57
2.5	The three support vectors for the final best line	57
2.6	Classification of Swarm algorithms optimisation	63
2.7	Salp’s series	68
3.1	Research design	82
3.2	Citrus tree treatment in the hydroponic system	93
3.3	Sample data for each class	94
3.4	Image pre-processing process	96
3.5	Image segmentation process	103
3.6	Colour-texture extraction process	107
3.7	SVM classification process	115
3.8	ANN process	119
3.9	SSA algorithm flowchart	124
3.10	The flowchart for proposed hybrid SSA-ANN methods	128
3.11	The flowchart for the proposed hybrid SVM-SSA	132
3.12	Independent validation datasets	137
3.13	Data diagram for each research objective	139
3.14	Validation process using Co-leaf datasets	141

4.1	Experimental design flowchart	147
4.2	Original image of each deficiency category	150
4.3	Percentage average for image size reduction for each class	151
4.4	Image (a) Before and (b) After lossless compression	153
4.5	Sample image (a) Before and (b) After contrast stretching	154
4.6	Distribution of PSNR after compression and contrast stretching	155
4.7	Distribution of structural similarity after compression and contrast stretching	158
4.8	The relationship between PSNR and SSIM	159
4.9	Histogram visualization of the image samples	160
4.10	Image channels and greyscale colour spaces	161
4.11	Result of Canny edge thresholding process	162
4.12	Different categories on Canny segmented	163
4.13	Sauvola thresholding process	163
4.14	Preview for each deficiency category on Sauvola segmented	164
4.15	Visual sample comparison between Canny vs Sauvola	165
4.16	Sample equalized image	167
4.17	Feature analysis Shapiro ranking of column features	170
4.18	Feature analysis of Pearson correlation	171
4.19	Feature analysis of mutual info classification	172
4.20	Average of variance for each class	173
4.21	Average of ASM for each class	174
4.22	Average of energy for each class	175
4.23	Graphic for 5-fold validation from Canny data	179
4.24	Training accuracy curves with cross-validation from Canny data	180

4.25	Training loss curves with cross-validation from Canny data	180
4.26	ANN model architecture	182
4.27	Training accuracy curves with Canny data	182
4.28	Training loss curves with Canny data	183
4.29	Confusion matrix from Canny segmented image using ANN	187
4.30	Class Prediction Error (CPE) from Canny segmented image using ANN	188
4.31	Graphic for 5-fold validation from Sauvola data	190
4.32	Training accuracy curves with cross-validation from Sauvola data	191
4.33	Training loss curves with cross-validation from Sauvola data	192
4.34	ANN model for Sauvola segmented data	192
4.35	Training accuracy curves from Sauvola data	193
4.36	Training loss curves from Sauvola data	193
4.37	Confusion matrix from Sauvola segmented image using ANN	196
4.38	Class prediction error from Sauvola segmented image using ANN	197
4.39	Graph of 5-fold validation for SVM-Canny data	199
4.40	SVM confusion matrix from Canny segmented image	203
4.41	SVM class prediction error from Canny segmented image	203
4.42	Graph of 5-fold validation for SVM-Canny data	205
4.43	Confusion matrix from Sauvola segmented image using SVM	209
4.44	Class prediction error from Sauvola segmented image using SVM	209
4.45	After SSA - ANN best model architecture from Canny segmented image	213
4.46	Convergence plot of ANN-SSA optimisation (from Canny segmented data)	214

4.47 Training accuracy curves with Canny data 215

4.48 Training loss curves with Canny data 215

4.49 SSA-ANN confusion matrix from Canny segmented image 220

4.50 ANN-SSA class prediction error from Canny segmented images 221

4.51 Convergence plot of ANN-SSA optimisation (from Sauvola) 223

4.52 After SSA - ANN best model architecture from Sauvola segmented image 224

4.53 Training accuracy curves with Sauvola data 224

4.54 Training accuracy curves with Sauvola data 225

4.55 ANN-SSA confusion matrix from Sauvola segmented image 229

4.56 ANN-SSA class prediction error from Sauvola segmented image 229

4.57 Convergence plot of SVM-SSA optimisation (from Canny) 232

4.58 SVM-SSA confusion matrix from Canny segmented image 238

4.59 SVM-SSA class prediction error from Canny segmented data 238

4.60 Convergence plot of SVM-SSA optimisation (from Sauvola) 240

4.61 SSA- SVM confusion matrix from Sauvola segmented image 243

4.62 SSA- SVM class prediction error from Sauvola segmented image 244

4.63 The comparison in accuracy score for each data 250

4.64 Results of the image enhancement 253

4.65 The results of image segmentation 253

4.66 Final prediction of unseen data 254

4.67 Image of a leaf that has mismatched judgments between experts and the system 259

4.68 Preview sample image after lossless compression 261

4.69 Visual sample after Sauvola segmentation 261

4.70 Confusion matrix for Co-leaf dataset using the proposed model 262

LIST OF ABBREVIATION

ANN	Artificial Neural Network
ASM	Angular Second Moments
AUC ROC	Area Under the Curve of the Receiver Operating Characteristic
CPE	Class Prediction Error
GLCM	Gray Level Co-Occurrence Metrix
HSV	Hue Saturation Value
PSNR	Peak Signal-To-Noise Ratio
RGB	Red Green Blue
SVM	Support Vector Machine
SSA	Salp Swarm Algorithm
SSIM	Structural Similarity Index Measure

CHAPTER 1

INTRODUCTION

This chapter discussed the background of the study, the problem statement, objectives, research questions, justifications and significance of the study. The contributions are separated into three parts: theoretical, practical and methodology. The end of this chapter discussed the scope of work and limitations of the research.

1.0 Background of the Study

Citrus fruits were the world's second most-produced fruit in 2021, with 161.8 million tonnes produced on over 10.2 million hectares (Gonzatto & Santos, 2023). Based on the data (World Citrus Organisation, 2022), worldwide orange production reached 76.7 million tons in 2021. Brazil, India, China, the United States, Mexico, Spain, Egypt and Indonesia are the central producer countries. Oranges are the most crucial citrus crop, accounting for 75.57 million tonnes (46.7% of total citrus fruit output) over a harvested area of 9.93 million hectares. Tangerines are the second most significant fruits, with a harvest area of 3.11 million hectares and an output of 41.95 million tonnes (25.9% of citrus fruit production). Lime and lemon output were 20.83 million tonnes (12.87% of citrus fruit production) on 1.34 million hectares (FAO, 2019).

There are four prominent citrus varietal groups distinguished in the international market (Singh et al., 2021):

1. Sweet orange (Citrus sinensis)
2. Lemon (Citrus lemon) and lime (Citrus aurantifolia)
3. Mandarins (Citrus reticulata)
4. Grapefruit (Citrus paradisi)

Mandarins, which include tangors, tangerines, clementines, satsumas, and tangelos, are the second-largest group of cultivated citrus after sweet oranges and account for approximately 25% of worldwide citrus production (Ladaniya, 2008). Mandarin refers to a group of citruses characterized by thin skin and easy peeling. Tangerines are a type of mandarin with a dark orange to reddish-orange colour and are smaller than Citrus reticulata (Usman & Fatima, 2018). The appearance of citrus fruit, as shown in Figure 1.1, is attractive if it is uniform in size, has a bright yellow-orange colour, is evenly and smoothly free from physical defects or previous pest attacks, and has a good taste and consistent sweetness.

Figure 1.1: Citrus reticulata Sp

The quality of citrus fruit is determined by its appearance and taste. The quality of citrus fruits depends on several factors, including environmental conditions, nutrient intake, planting method, and treatment at and after harvest. Vitamins are important factors for a plant's growth and reproduction. Macronutrients, which include phosphorus (P), nitrogen (N), potassium (k), sulfur (S), calcium (Ca), and magnesium (Mg), are regularly used to perceive plant nutrients. Nutrient consumption is a massive aspect of cultivating citrus (Sun et al., 2018) (Jeyalakshmi & Radha, 2017). Micronutrients, which include zinc (Zn), boron (B), sodium (Na), copper (Cu), nickel (Ni), iron (Fe), chlorine (Cl), cobalt (Co), silicon (Si), molybdenum (Mo), and manganese (Mn), are also needed by plants.

Plant nutrient deficiency symptoms are usually visible in leaves and fruits. Symptoms in leaves include marginal chlorosis, interveinal chlorosis, uniform chlorosis, distorted edges, reduction in leaf size, and necrosis. Although similar symptoms can occur in old and new leaves, the deficient nutrient may vary depending on other factors. Deficiency symptoms are widely used to identify nutrient responsiveness in citrus leaves (Jeyalakshmi & Radha, 2017).

Research on identifying nutrient deficiencies in plants using image processing techniques, especially in Indonesia, has not been conducted extensively, as most research focuses on identifying plant diseases. As an agrarian country, Indonesia must develop the latest agriculture technology. Nutrient deficiency in plants can result in decreased crop yields. Research on nutrient deficiencies has been conducted (Rahadiyan et al., 2023), but the research only had an accuracy of 59% because it only classified leaves based on texture. Other studies recommend a system equipped with an automatic identification system for plants (Jose et al., 2020).

Action must be taken to correct nutritional shortages in plants and reduce losses. Numerous auxiliary technologies have been created to automate information and data obtained from picture processing. This technology's design and use will greatly help to lower expenses, increase productivity, and decrease the need for chemical fertilisers. Plant nutrition detection is a new and popular research topic, but many challenges exist in applying image processing expertise and learning procedures for plant nutrition recognition. The suggested algorithms must be precise with a small error margin since inaccurate detection can seriously impair agricultural output.

The topic of intelligent classification systems study encompasses a wide range of activities, including feature extraction, segmentation, preprocessing of data, and classifier learning. Image identification is one of the most interesting challenges at hand. Techniques for image processing are essential for analysing smart agricultural technology. A detection and classification system using digital image processing technology would more accurately and quickly classify nutrient deficiency symptoms in citrus trees than a human eye could (Tian et al., 2020). It would enable farmers to take appropriate corrective action in farm management. The primary goal of this proposed research is to develop detection and classification methods for nutrient deficiencies in citrus leaves using image analysis techniques. The purpose is to perceive and quantify the colour and texture attributes of citrus leaves the use of the following steps (Sun, 2016):

a. Image pre-processing
b. Image segmentation (based on threshold and edge technique).
c. Feature extraction (texture and colour extraction)
d. Image classification using ANN and SVM
e. Improved ANN and SVM using the Salp Swarm Algorithm (SSA)

In this modern era of technology, advancements in artificial intelligence have given rise to two primary paradigms in data analysis and understanding: machine learning and deep learning. As a subset of artificial intelligence, deep learning has garnered significant attention for its ability to tackle complex problems such as image recognition, natural language processing, and various other tasks. Deep learning has achieved remarkable feats across domains by utilising deep neural network architectures.

However, despite the remarkable popularity and achievements of deep learning, the older machine learning technology still holds a firm place in data analysis. Machine learning has proven effective in solving classification and prediction problems with lower resource requirements and higher interpretability. It is crucial to acknowledge that while deep learning can deliver exceptional results in complex scenarios, there are situations where the more established the machine learning approach is, the more suitable it is. In some cases, limited dataset sizes or more superficial problem characteristics might not necessitate the level of complexity offered by deep learning. The research has a relatively small dataset (1200 data), and complex deep learning models might be complicated to train effectively and could suffer from overfitting. In such cases, simpler machine learning models might be more suitable. Complex deep-learning models often require extended training and significant computational resources (Lai, 2019). The research only needs moderate accuracy, and machine learning models can provide satisfactory results. Several crucial considerations come into play when utilizing Deep Learning to process relatively small datasets. The foremost risk is overfitting, as complex Deep Learning models might memorize training data instead of grasping general patterns, leading to poor performance on new or real-world data. Addressing overfitting demands techniques such as regularization, dropout, and feature dimension reduction.

Additionally, limited dataset representation hampers the model's ability to discern intricate patterns, restricting its generalization capacity. Striking the right balance of model complexity is essential, as overly complex models can exacerbate overfitting issues on small datasets. Moreover, Deep Learning models' substantial computational demands, especially for deep neural networks or large models, might not yield proportional benefits on small datasets (Wang et al., 2021). Lastly, data augmentation proves vital; introducing variations through simple transformations like rotations or cropping bolsters the training dataset's diversity, curbing overfitting risks and bolstering generalization capabilities. Because this system needs to be adopted by agriculture, especially farmers, choosing machine learning, which is easier to interpret results and requires less memory, will be better than using deep learning.

The application of image processing technology and the development of algorithms for identifying plant nutrition presented challenges. Large labelled datasets cannot be created because labelling pictures for training models is costly, time-consuming, and laborious (Meir et al., 2023). However, using electroencephalograms from plant pathology experts can significantly reduce the labelling time while incorporating expert knowledge into artificial intelligence models. Another challenge is identifying nutritional inadequacies in crops, which can be addressed through image processing and Convolutional Neural Networks (CNN) (Mishra et al., 2023). CNN has shown high accuracy in detecting and diagnosing nutrient deficits in different cultivars (Raju & Narasimhaiah, 2023). Additionally, there are various approaches and techniques in image processing for identifying disease conditions in plant leaves, but the optimal approach is yet to be determined (Baraskar et al., 2023). These challenges highlight the need for further research and development in image processing and learning algorithms for plant nutrition recognition.

Various classification algorithms have been used for leaves. These include K-Nearest Neighbor Classifier (KNN), Probabilistic Neural Network (PNN), Genetic Algorithm, Support Vector Machine (SVM), Principal Component Analysis, Artificial Neural Network (ANN), Fuzzy Logic, Decision Tree (DT), Naive Bayes (NB), Radial Basis Function Neural Network, Random Forest, and Self Organizing Map (Bhosale et al., 2023) (Venkatakrishnan & Natarajan, 2023) (Aslan, 2023) (Mangaoang &