International Journal of Engineering and Science Invention (IJESI)
ISSN (Online): 2319-6734, ISSN (Print): 2319-6726
www.ijesi.org ||Volume 13 Issue 9 September 2024 || PP 117-126

Analysis of Coconut Datasets for Machine Learning and Deep Learning Models

Shankara. N.B[1], Prasadu Peddi[2]
[1]*Research Scholar, Dep of CSE, Shri Jagdishprasad Jhabarmal Tibrewala University Jhunjhunu, Rajasthan, India*
[2]*Professor, Dep of CSE & IT, Shri Jagdishprasad Jhabarmal Tibrewala University Jhunjhunu, Rajasthan, India*

ABSTRACT
This paper discusses the development of an automated disease detection system for coconut plant monitoring, focusing on the integration of multiple machine learning techniques, real-time detection capabilities, scalability, and adaptive learning. Deep learning models, particularly Convolutional Neural Networks (CNNs), can autonomously derive image properties associated with illness symptoms. The selection of appropriate architectures, such as ResNet, VGG-16, or EfficientNet, facilitates the capture of complex patterns within the data. The study examines the feasibility of identifying and assessing the health of coconut trees using high-resolution images combined with deep learning approaches. The Resnet-50 model outperforms the VGG-16 architecture in detection and health classification tasks, showing that most affected coconut trees have Ganoderma infections and potassium deficiencies. The proposed method shows potential for coconut tree management in Thailand, allowing for more efficient use of workers and less time spent in the field. To maximize model performance, future research should aim to increase the quantity and diversity of datasets, including various visual attributes. To better classify health problems, future studies may use a multi-spectral camera. By combining supervised, unsupervised, and semi-supervised learning approaches, the system can be tailored to the complexities of coconut plant monitoring and broader applications.
KEYWORDS: *Coconut Leaf, ResNet, VGG-16 and CNN.*

I. INTRODUCTION

Machine learning plays a crucial role in disease detection in agriculture, particularly in identifying diseases in coconut data. The classification of diseases depends on the characteristics and research setting of the data point. Contextual diseases are observed when seasonal data changes outside the anticipated time range, while manual disease detection techniques involve dashboard evaluations and monitoring for unusual fluctuations. Design factors for machine learning-based automated sickness detection systems include event definition, scalability, brevity, rate of change, and timeliness. The importance of finding coconut diseases affects the choice of machine learning algorithms and the tasks and goals the algorithms are meant to achieve. Machine learning disease detection systems have three ways to achieve conciseness: univariate, multivariate, and mixture methods. Choosing the incident description format is critical as illness diagnosing systems are not entirely automated. Supervised learning-based disease detection is not effective in finding novel diseases and anomalies, while unsupervised learning gradually acquires characteristics of typical behavior and recognizes abnormalities when the examined data deviates from the designated model.

Machine learning and deep learning are essential tools for automating intrusion and disease detection in coconut palms. Understanding normal and abnormal behavior is crucial for automating intrusion and disease detection systems, as they often receive large, highly variable, and non-linear data inputs. Machine learning algorithms like ML and DL constantly update the new normal using input data, making it important to minimize the relevance of unusual data points to prevent incorrect conclusions in the future.

Neural networks (NNs) are indispensable for disease detection in coconut palms and natural language processing. Data normalization, augmentation, and optimization algorithms like RMSprop, Adam, and Stochastic Gradient Descent are needed for deep learning (DL). Transfer learning, GPUs, and TPUs accelerate deep learning model training, while cloud computing and distributed training can boost scalability.

AI and DL models use neural networks to improve accountability and transparency, while multi-modal learning enhances model understanding and effectiveness. Legal, ethical, and legislative frameworks are needed to address issues like unemployment, discrimination, and resource inequality. Online learning and meta-learning help models learn quickly and adapt to new situations without past knowledge. Continuous learning is crucial for AI systems to adapt to new challenges and retain past information. Advances in quantum computing are revolutionizing computational approaches, such as customised capsules, transformer models, and Kaps Net. Auto

ML automates model selection, hyperparameter optimization, and architectural exploration, while specialized neural networks manage social network and recommendation system graph topologies. Deep learning (DL) is a subset of machine learning that prioritizes input comprehension and integration, with multi-modal learning integrating text, visual, and audio information. Deep reinforcement learning (DRL) uses explainable artificial intelligence (XAI) to explain agent behavior, particularly in safety-critical fields like healthcare and autonomous robots. Emotional AI models assess coconut tree disease and textual sentiment in affective computing, enabling emotionally intelligent virtual assistants, empathy-enhanced AI systems, and mental health assessments. Self-supervised learning methods improve Natural Language Processing (NLP) models, and evolutionary neural network training optimizes structures and parameters using genetic variation and natural selection. Differentiable rendering uses neural networks and rendering engines to create 3D visuals from 2D photos. Deep learning has revolutionized various industries, including computer vision, robotics, augmented reality, quantum chemistry, and climate prediction. It has solved mathematically complex quantum mechanical problems, predicted chemical characteristics, found novel materials, and accelerated medicine development. Transfer learning reduces linguistic barriers by using pre-trained models in resource-abundant languages to improve performance in languages with less data. Deep neural networks have been used to investigate, predict, and model climate data, improving our understanding and allowing more accurate predictions and policy decisions.

Hartrot, a wilt disease, is a perishable fruit disease that can be deadly if left untreated. It affects coconut trees in countries like Brazil, Ecuador, Guyana, Peru, Suriname, Venezuela, Nicaragua, Costa Rica, Honduras, and the West Indies. The disease is linked to the systematic colonization of the plant by a non-pathogenic protozoan belonging to the Trypanosomiases family. The Phyto monas genus infects plants that do not produce latex but may damage fruit. The disease's outward manifestation is curling around the margin's lower leaves, burning and drying out leaflets. As palm trees mature, their trunks taper sharply towards the base crown, and inflorescent shrinks in number and size. The newly-emerging leaves are noticeably smaller, and the palms stop producing fruit when their crowns shrink and eventually separate from the stem. The decay of the trunk caused by termites, such as Marasmiellus cocophilus, is a devastating disease affecting coconut trees in East Africa.

II. METODOLOGY

Deadly yellowing (LY) is a disease that has been affecting coconut trees in western Jamaica since the 1800s, and it is now distributed to southern Mexico, Belize, Honduras, Cuba, the Dominican Republic, Haiti, Jamaica, Cayman Islands, and the Bahamas. The worldwide supply chain for coconuts is in grave danger due to the fast proliferation and subsequent extinction of palm species and the lack of a practical solution. Nearly two-thirds of the world's coconut palms are in risk of extinction, according to Harris (1978). The most noticeable sign in fully grown, tall-type coconut palms—the most susceptible—is the premature dropping of most or all fruit. Injured fruits often have a brownish-black colour and a wet appearance at the calyx end. The blossom stems of healthy palms may be any shade from milky white to yellow. As they emerge from the surrounding spathe, flower stalks with inflorescence necrosis show partial or complete blackening. The lower, older leaves show the first signs of foliar yellowing, which eventually affects the whole crown. The first indicator of this disease might be a single golden leaf, called the flag leaf, located in the middle of the crown. The yellowed leaves keep their turgidity and become brown; they stay this way for a long time until they dry up, wilt on the stem and fall off. The bud's demise is signaled by the decay of leaf tissue caused by secondary bacterial invasion and the newly created spear leaf's structural failure, which is shown by its drooping from the crown. After three to six months of the initial symptoms appearing, the palms usually start to decrease due to the progression of LY. Dwarf coconuts show signs such as early nut and flower abscission, a noticeable browning of mature leaves, and the usual yellowing that tall-type coconuts experience. At times, the 'Malayan Green Dwarf' will show symptoms of wilting on the leaves, most notably inflorescent that wither away inside their closed spathes without ever emerging. Significant flaccidity and inward folding are seen in the pinnae of mid- and upper-crown leaves, and as the sickness progresses, constrictions in the stem may also become visible. The study presents a new operational model for the biophysical classification of coconut trees using microwave remote sensing and a theoretical framework. Building a theoretical model that accurately represents coconut trees is the first step in studying their back scattering mechanism. Before conducting any investigation, it is essential to gather enough ground truth data, particularly for purposes of categorizations and theoretical validation. Using biophysical elements as inputs to a theoretical model, which will then create simulated outputs, we can get a better knowledge of how microwaves interact with the structural composition of coconut trees. Modeling simulation results may provide light on the connection between the many biophysical coconut tree properties and the various derivatives of the scattering matrix, such as the back-scatter coefficient. Machine learning classification criteria will be enhanced by a better understanding gained from this data.

In order to improve classification alongside the biophysical data, it is essential to get the GPS coordinates of the coconut trees that were surveyed. To evaluate the classification system's performance, it is essential to establish ground truth data. Getting the plantation owner's agreement is the first step in surveying varied places of interest.

The work of Teng et al. (2015) and Koay et al. (2008) will be the basis for developing the theoretical model. The radiative transfer theory serves as the foundation for the model, as shown in Figure 1. The canopy of coconut trees is represented as a two-layered, flat-surfaced stochastic medium. The pinnae of coconuts are shown as oval discs with cylindrical stems and fronds. Beneath the outermost layer composed of pinnae and fronds is the coconut stem. In practical applications, the Integral Equation Model (IEM) is used to analyze ground-based back-scatter phenomena.

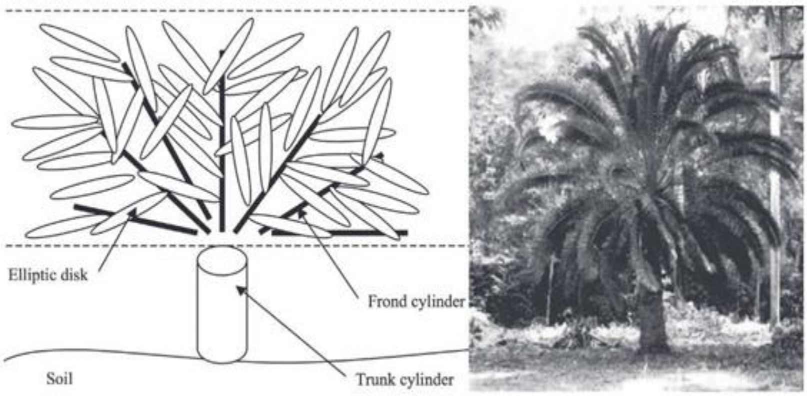


Figure 1: coconut tree illustration of the Model.

Based on reviews and our own requirements, we chose MATLAB as the programming platform to develop the deep learning model. This is because we can also use MATLAB to analyze data and also use MATLAB classical machine learning techniques such as SVM to do classification. Being on the same platform allows us the ease of comparing data and performance of each classification model.

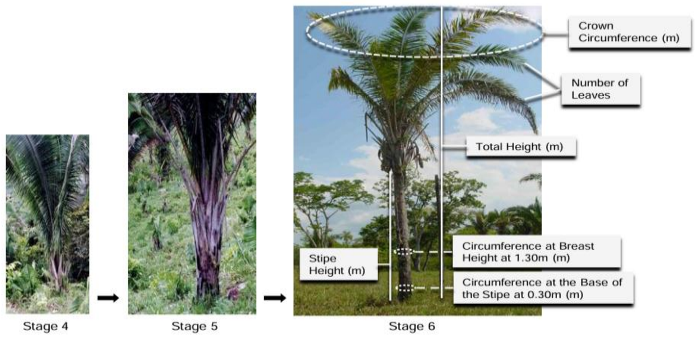


Fig. 2. Benfica babassu coconut tree life-cycle phases and structural data

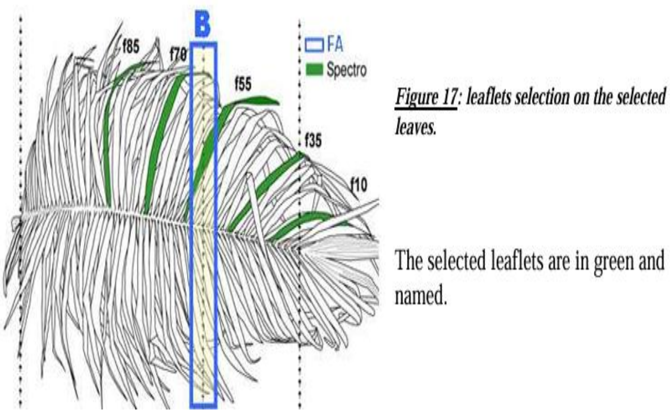


Fig: 3 cutting leaves, spectral acquisitions are performed on a number of chosen leaflets:

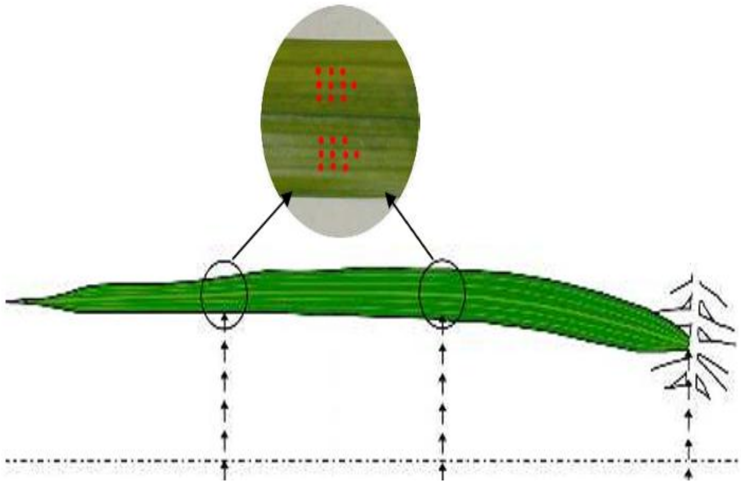


Fig 4 Reflectance spectral acquisitions' protocol on coconut-trees' leaflets

Researchers used reflectance spectrum analysis to categorize coconut trees based on severity of Ganoderma disease using spectroscopy and remote sensing technologies. The study divided the trees into three or four categories based on their severity: healthy, medium, low, and high. Visual diagnostic assessments were included in the database with reflectance spectroscopic data from the chosen coconut trees. The previous approach lacked statistical reliability due to the small sample size used in the study. The present project uses standardized statistical analysis techniques to collect a large data set, with plans for an even higher sample size. Remote sensing techniques can help monitor Ganoderma's growth, potentially leading to the creation of control techniques or therapies. Foliar diagnostics is the most effective strategy for detecting Ganoderma infections by examining the presence of mycelium in coconuts. The study also considers the B point as an important reference point for establishing foliage features. Weather and cloud thickness affect the acquisition process, and image processing and segmentation are crucial digital functionality for reliable and precise results.

Modern plant disease detection and classification systems often miss important details like texture and features, leading to subpar photos. The AFKMRG methodology produces higher quality and richer information sets than K-Means or Region Growing techniques alone. Global threshold methods struggle to handle localized intensity changes and changing lighting circumstances. Adaptive threshold and FCM anomaly segmentation improve picture quality.

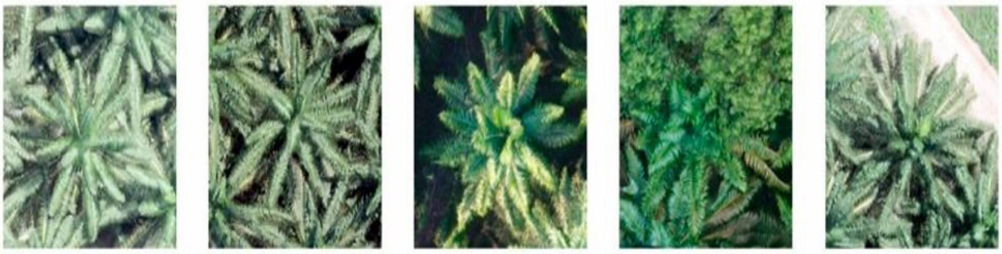
Accurate plant picture re-processing and segmentation are crucial for disease detection, and deep learning and machine learning are being used to mechanize these processes. Optimization techniques, such as particle swarm optimizations and differential evolution, can make disease detection systems more accurate and efficient.

Advanced optical monitoring of plant leaves can improve disease diagnosis efficiency and facilitate seamless integration with treatment optimization tools. Leaf image preparation is essential for accurate feature extraction, rapid disease diagnosis, and precise plant categorizations. Supervised learning and unsupervised learning algorithms are used to assign pixels based on user-provided labels, making the process more efficient and reducing the likelihood of mistakes.

Transforming shortstops from. TIF to.JPG was the first stage in processing UAV images with RGB bands. Memory and processing power consumption was greatly impacted by the image sizes used for testing and training. As a result, we reduced the size of each image while keeping the original proportions. Consequently, to avoid any possibility of overlap, each image was adjusted to fit into sections of 2000 by 2000 pixels.

In line with the development of training datasets for deep learning-based automated identification and measurement of coconut trees.

A robust collection of coconut trees and a weaker set of coconut plants have been meticulously separated in the training data. A section of the training data set was devoted to healthy coconut trees, which were photographed in a variety of lighting and environmental settings and included both high- and low-resolution images. Making a training data set and developing a classification model for coconut trees. In order to automate process of counting and identifying coconut trees using deep learning algorithms, this solution makes use of re-classified data. There are now two separate databases, one including perfectly healthy coconut palms and other including very unhealthy ones. In order to better identify healthy specimens, the training data set included a varied variety of pictures of coconut trees in different states. This includes haze effect scenarios and partial canopy circumstances, both of which depend on preserving at least half of the canopy.



(a) Healthy coconut tree tree

(b) Unhealthy coconut tree tree

Figure. 5 Examples from the (a) healthy and (b) unhealthy datasets.

The study aimed to establish a training set of healthy and unhealthy coconut trees by separating each tree from other relevant factors. The training data set included coordinates for every canopy border, bounding boxes around each canopy, and a dataset.xml file was prepared using the Pascal VOC dataset guidelines. A testing data set was also created, with 10% of the total data set reserved for testing. The primary network architecture used in the experiment was Faster R-CNN, with two models: VGG-16 and ResNet-50. These models were built using Python 3's TensorFlow and Keras deep learning frameworks. The model training operation was conducted on a Windows 10 PC with a Ge Force RTX 2080 Ti. Training for the VGG-16 network was halted after 39 epochs and for the ResNet-50 network at 40 epochs. Dietary deficiencies are the leading cause of coconut tree health decrease, as essential nutrients such as nitrogen, phosphorus, potassium, magnesium, and boron may be applied to coconut trees to alter their physical characteristics. Coconut trees are now being affected by an epidemic. Researchers investigated the health of coconut trees in terms of color, crown density, and crown size. A yellow canopy indicates a nitrogen deficiency, while canopies of coconuts deficient in potassium appear greenish orange. Most coconut trees lacking in nutrients lacked many essential components, rather than just one. When trees experienced severe nutritional deficiencies, their crown density and size shrank.

A substantial increase of Ganoderma has had a negative impact on coconut trees, with lesions in crown size and density leading to dryness and abscission of leaves. This issue, known as Ganoderma, considerably slows down development and decreases output. Most coconut trees showed appropriate conditions in this study, according to an analysis of physical parameters found in UAV data and plot surveys. In contrast to the diseased specimens, the photogenic coconut trees showed much larger crown diameters, richer green hues, and thicker fronds.

A systematic framework for categorizing health based on outwardly apparent characteristics was born when all factors were considered together. Young coconut trees may still maintain perfect health and vigor, while coconut trees with large crown diameters and poor densities had negative outcomes. The study involved analyzing coconut tree diseases using high-resolution RGB images captured by UAVs. The images were divided into 2000 × 2000-pixel sections for optimal processing efficiency and memory accessibility. The data set consisted of 133 images, 17 for testing and 116 for training. The results showed that coconut trees with yellowing canopy were deficient in nitrogen, potassium, and other nutrients. Crown density and size were reduced due to severe nutritional deficiencies. During the field inspection and investigation phase, it was found that coconut trees were affected by an advanced stage of Ganoderma, which leads to dryness and abscission of leaves. A deep learning model was trained and assessed using the same data set for Resnet-50 and VGG-16, two fundamental models built on the Faster-RCNN architecture. The model's efficacy and capabilities were evaluated throughout training and testing. The research used a variety of setups and heights to evaluate the autonomous coconut tree identification. UAVs were deployed at 50, 80, and 90 meters of height, but coconut palms, common in low-altitude photos, could be difficult to mosaic. To improve illumination, operations were carried out at different intervals and used different color profiles. A D-log colour profile between 10:00 AM and 3:00 PM was found to be suitable, according to the test findings.

The combination of UAVs and sophisticated CNN architectures to detect and count coconut tree diseases has the potential to dramatically improve agricultural surveillance. Addressing the difficulties highlighted in the study will enhance its relevance in the area.

Image	**Resnet-50**			**VGG-16**		
	Coconut Tree	Healthy	Unhealthy	Coconut Tree	Healthy	Unhealthy
1	42	42	0	40	40	0
2	41	41	0	41	41	0
3	39	39	0	36	36	0
4	13	13	0	13	13	0
5	6	6	0	6	6	0
6	18	18	0	18	18	0
7	21	21	0	21	21	0
8	28	24	4	28	28	4
9	36	36	0	36	34	2
10	47	25	22	47	24	23
11	53	50	3	53	50	3
12	65	45	20	65	55	10
13	22	20	2	22	21	1
14	32	30	2	32	30	2
15	43	40	3	43	40	3
16	28	20	8	28	22	6
17	32	29	3	32	28	4
18	45	35	10	45	35	10
19	39	34	5	39	35	4
20	80	78	2	80	78	2

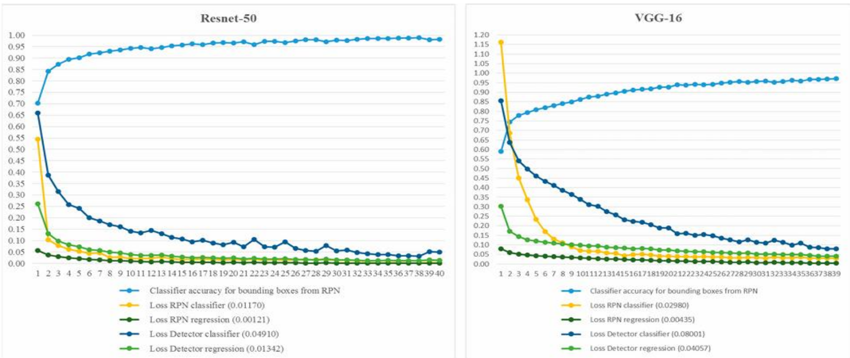


Figure.6 Comparison of model training performance.

While evaluating the model's performance in real-time, you can see its accuracy and loss metrics. As the research shows, the Resnet-50 network outperforms the VGG-16 network by a significant margin. As can be seen from the pattern, the two models' values have comparable tendencies in both rising and falling trends. Improving model selection performance may be achieved by evaluating it using many metrics.

A comprehensive analysis of the methods used by various models to control the duration of each period. The graph shows that compared to VGG-16, Resnet-50's processing time was around 10 minutes shorter. There was a time difference of around 5 hours; ResNet-50 required about 40 hours to analyse the data, whereas VGG-16 required about 45 hours.

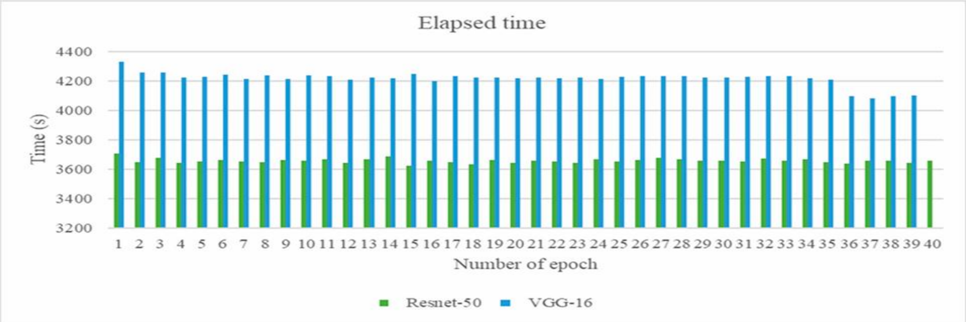


Figure 7 Processing time.

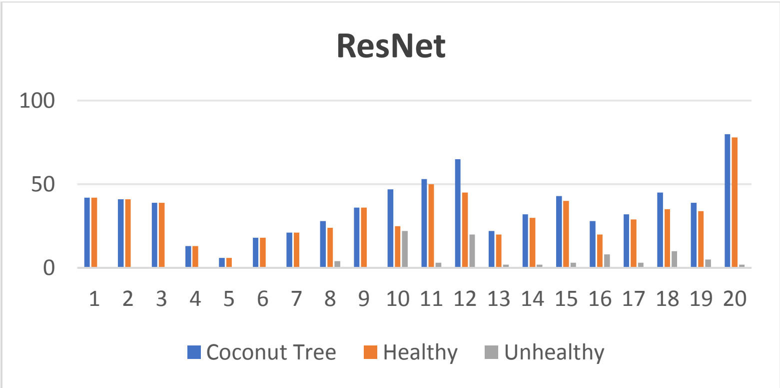


Figure :8 Rest Net classification on Over all Data

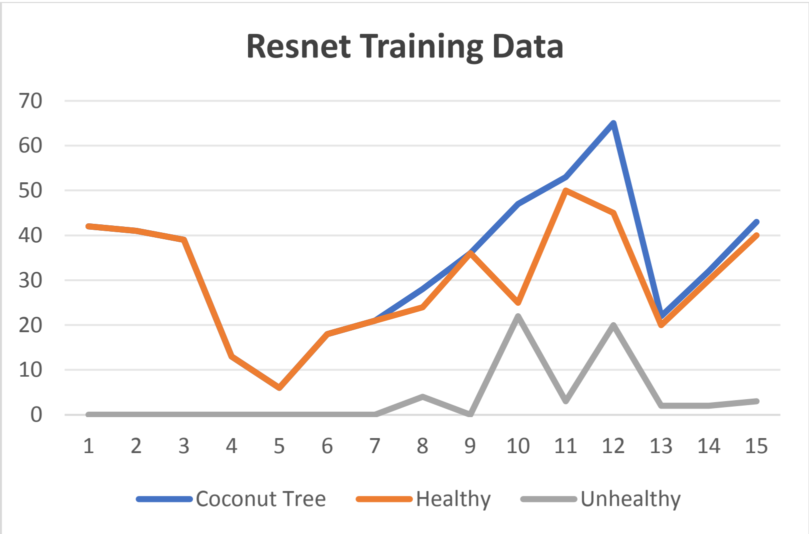


Figure :9 Res Net Training Data

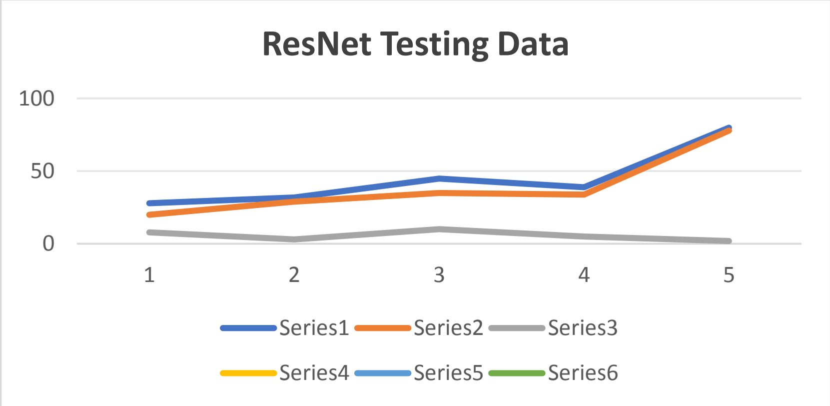


Figure 10 ResNet Testing Data

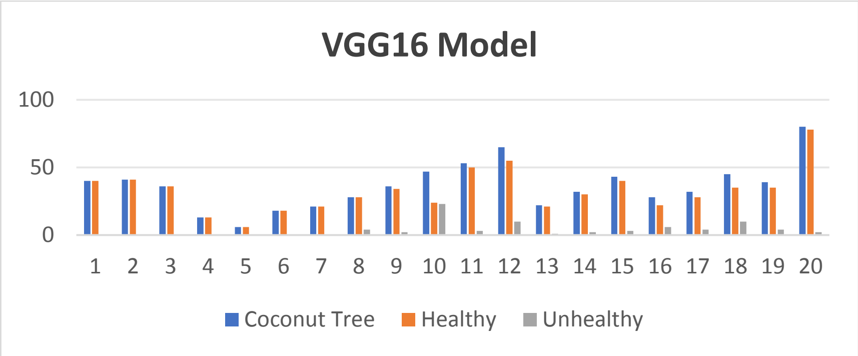


Figure :11 VGG 16 Model

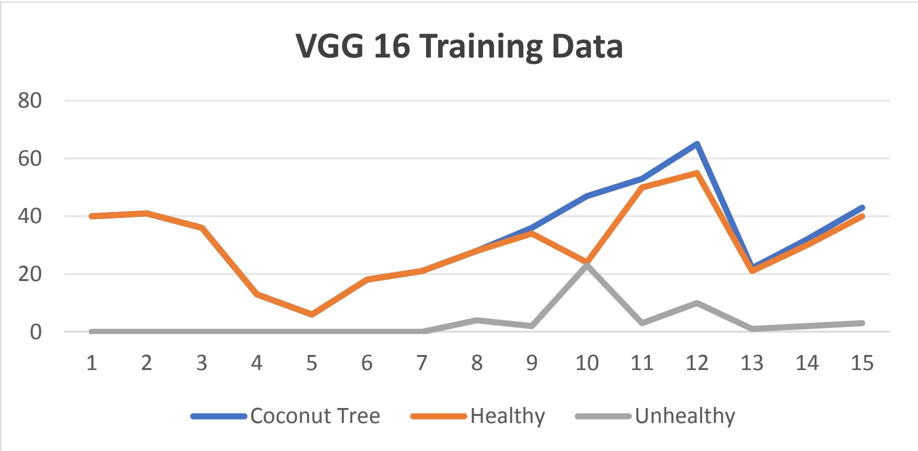


Figure 12 VGG 16 Model Training Data

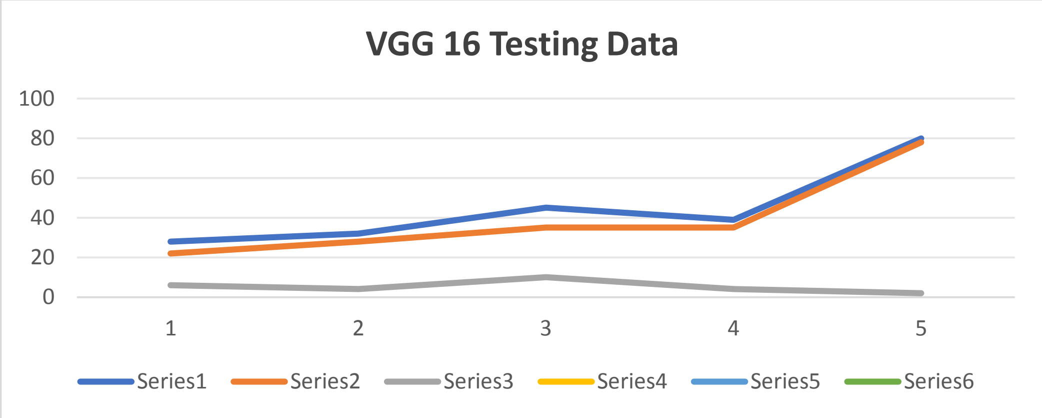


Figure :13 VGG model Testing Data

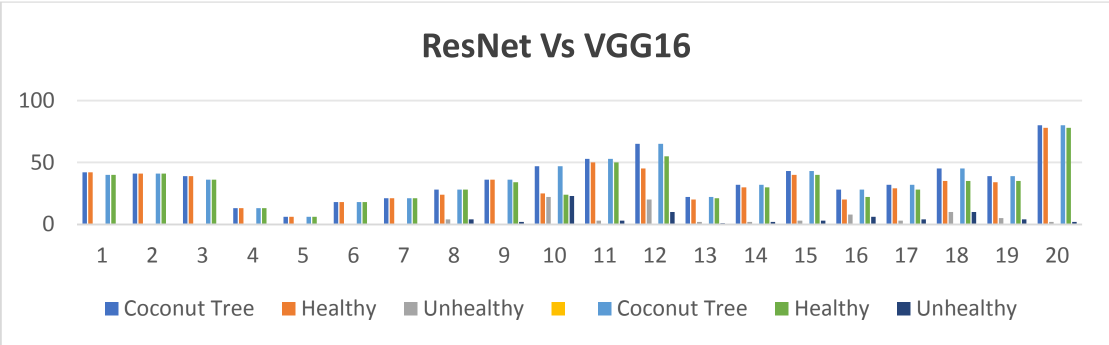


Figure 14 Res Net Vs VGG16 Data

III. CONCLUSION

The Resnet-50 model was found to be the superior choice for evaluating the health status of coconut trees, with an accuracy rate of 91.24%. The model was used in a comparative analysis of 21 samples from coconut palms, revealing that all 223 coconut plants were in optimal condition, while 6 exhibited signs of disease. The predictive capabilities of the model were optimized for coconut trees, identifying instances of classifications regarding damaged coconut plants. The analysis revealed potential health issues in six out of fifteen coconut plants, including nitrogen deficiencies, potassium deficiencies, and boron deficiencies. The study found that nitrogen deficiencies are the primary factor contributing to health issues observed in coconut plants, with the model accurately predicting that 50% of infected coconut trees will sustain damage. The study highlights the challenges in coconut farming and emphasizes the potential of predictive models to improve the health of coconut populations and encourage sustainable farming by focusing on combined management strategies and early detection.

REFERENCES

[1]. Ananda S. Paymode and Vandana B. Malode, Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG, Artificial Intelligence in Agriculture, Vol. 6, pp. 23-33, 2022.

[2]. Chaki, J., Dey, N., Moraru, L., Shid, F., (2019). Fragmented plant leaf recognition: Bag-of-features, fuzzy-color and edge-texture histogram descriptors with multilayer perceptron, Optik, 181, 639-650

[3]. G. J. Jagtap, P. Peddi and Y. D. Sinkar, "Application of YoloV5 for Brain Tumor Detection from MRI," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024, pp. 1-5, doi: 10.1109/ESCI59607.2024.10497232.

[4]. Kaur, S., Pandey, S., Goel, S.: An automatic leaf disease detection system for legume species, Journal of Biological Today's World, 6, 115-122 (2017).

[5]. Jyotismita Cand Nilanjan D. (2018). A Beginner's Guide to Image Pre-processing Techniques, CRC Press - Series: Intelligent Signal Processing and Data Analysis.

[6]. Sachdeva, G., Singh, P., Kaur, P. (2021). Plant leaf disease classification using deep Convolutional neural network with Bayesian learning, Materials Today: Proceedings, 45(6), 5584-5590.

[7]. Rajesh, B., M. Vishnu Sai Vardhan, and L. Sujihelen. Leaf disease detection and classification by decision tree. 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184). IEEE, 2020.

[8]. Pethybridge, S. J., & Nelson, S. C. (2015). Leaf doctor, a new portable application for quantifying plant disease severity, Plant Disease, 99(10), 1310-1316

[9]. Prasadu Peddi, & Dr. Akash Saxena. (2015). The Adoption of a Big Data and Extensive Multi-Labled Gradient Boosting System for Student Activity Analysis. International Journal of All Research Education and Scientific Methods (IJARESM), 3(7), 68-73.

[10]. T. N. Pham, L. V. Tran, and S. V. T. Dao, Early Disease Classification of Mango Leaves Using Feed-Forward Neural Network and Hybrid Metaheuristic Feature Selection, IEEE Access, Vol. 8, pp. 189960-189973, 2020.

[11]. Yun Zhao, Cheng Sun, Xing Xu and Jiagui Chen, RIC-Net: A plant disease classification model based on the fusion of Inception and residual structure and embedded attention mechanism, Computers and Electronics in Agriculture, Vol.193, pp. 106644, 2022.

[12]. M. Lv, G. Zhou, M. He, A. Chen, W. Zhang, and Y. Hu, ``Maize leaf disease identification based on feature enhancement and DMS-robust AlexNet," IEEE Access, Vol. 8, pp. 57952-57966, 2020.

[13]. Prasadu Peddi & Shankara. N.B. (2024). Applying deep learning and machine learning to the study of coconut palm disease. Internation Journal Of Advance Research And Innovative Ideas In Education, 10(4), 3394-3399.