J. Electrical Systems 20-4s (2024): 524-531

[1] Charlot L. Maramag
[2] Thelma D. Palaoag

Unlocking Machine Learning Algorithms for Bambooshoots.AI: Revolutionizing Agricultural Applications with Computer Science

Journal of Electrical Systems

Abstract: - This research initiated a comprehensive investigation to determine the most suitable Machine Learning (ML) algorithms for bamboshoots.AI application. Convolutional neural networks (CNNs) and support vector machines (SVMs) were cited as two essential ML techniques in a comprehensive review of ten papers. To further assess these two algorithms in the context of the Bambooshoots.ai application, an experimental analysis was also carried out. The CNN model had been trained over 25 epochs with 8 batches of data per epoch, demonstrating a consistent increase in accuracy, and reached 97.94% by the end of training. In contrast, the SVM model provided an accuracy of approximately 57.14%. The experimental results indicated that the CNN model had been better at classifying bamboo shoot images, thus making it the preferred choice for the BambooShoots.AI application. The significant difference in accuracy between the two models, as well as the consistent performance of CNNs in image-based tasks, informed this decision. The results of this research were thrilling because they showed that machine learning could be used to help understand and manage bamboo shoots. This could have been a big help to farmers and could have led to new ways of doing things in agriculture.

Keywords: Machine Learning, Accuracy, ML Algorithms, Bamboo Shoots, CNN, SVM, Computer Science.

I. Introduction

The advent of Machine Learning has been instrumental in advancing agricultural practices, fostering improvements in efficiency and precision agriculture [1]. They enable more accurate forecasting, improved resource management, and enhanced crop yield, driving sustainable and efficient agricultural practices[2]. There are numerous aspects of agriculture where machine learning (ML) is now being used. They have revolutionized crop yield prediction[1], optimized irrigation systems[3], facilitated precise identification of crops and weeds[4], and plant disease and pests' detection[5],[6],[7]. The efficacy of such applications significantly hinges on the appropriate selection of ML algorithms, with enhance their learning capabilities and improve their ability to adapt and yield accurate results.

According to [2], ML-enabled precision agriculture can offer benefits such as increased yield, reduced input cost, and improved environmental sustainability. Also, [8] discuss in his article that integrating ML can transform agriculture by fostering data-driven decisions and optimizing farming practices. Guided by these insights, our research aims to investigate diverse ML algorithms to assess their compatibility and efficacy within the BambooShoots.AI application.

The selection of an optimal Machine Learning (ML) algorithm for any AI application is crucial, profoundly impacting its performance, functionality, and utility. The choice of an ML algorithm is not just a technical consideration, but also a determinant of the AI application's ability to learn, adapt, and provide accurate insights from the processed data. This importance is further magnified in applications such as BambooShoots.AI which aims to determine the optimal harvesting height and diseases of bamboo shoots.

The motivation for this research stems from the need to optimize the BambooShoots.AI application, aligning it with the most suitable ML algorithms. Given the variety of ML algorithms available, each with its unique strengths, weaknesses, and applicability, identifying the right fit for BambooShoots.AI poses a substantial research opportunity. This exploration will not only enhance the efficiency and effectiveness of BambooShoots.AI but also contribute valuable insights to the broader discourse on ML algorithm selection for AI applications.

The research problem at hand involves identifying the most suitable Machine Learning algorithm(s) for BambooShoots.AI. Addressing this problem requires systematically exploring and evaluating various ML algorithms and understanding their applicability in the context of BambooShoots.AI.

This research is about more than just making BambooShoots.AI better. This study aims to find the most suitable machine learning (ML) algorithms that fit well with the specific needs and goals of the platform. Specifically, it aims to (1) collect and preprocess data for ML algorithm evaluation; (2) evaluate and compare the performance of different ML algorithms; and (3) make a data-driven decision based on the evaluation results.

[1, 2] College of Information Technology and Computer Science, University of the Cordilleras Baguio City, Philippines
* Corresponding Author Email: charlottalagutan@gmail.com, tdpalaoag@uc-bcf.edu.ph

In doing this, there is a hope to improve the understanding of how to select ML algorithms for AI applications in agriculture. The knowledge gained from this research is expected to enhance BambooShoots.AI's capabilities and provide essential advice for other similar AI platforms used in the agriculture sector. Essentially, this research is about promoting a new era of creativity and efficiency in agricultural practices through the careful use of AI and ML.

II. METHODOLOGY

A. *Comprehensive Literature Review*

The research began with a systematic literature review aimed at deepening the understanding of Machine Learning (ML) applications within the context of agricultural Artificial Intelligence (AI), especially concerning bamboo shoots. This review provided a comprehensive assessment of the existing literature, focusing on various ML algorithms' applications in AI.

The systematic review's primary goal was to discern the frequency and performance of different ML algorithms, as recorded in academic literature. The selection of ML algorithms for further experimental exploration was driven by two considerations: the prevalence of their use across diverse studies (as evidenced by their mention in numerous academic journals), and their documented accuracy within relevant AI contexts.

Subsequently, the ML algorithms that were most frequently used and demonstrated superior accuracy in agricultural AI applications, according to the extant literature, were selected for further experimental investigation in this study.

B. *Data Collection*

The initial research phase involved identifying and collecting pertinent datasets, with a particular focus on bamboo shoots data regarding growth rates and diseases. These datasets were obtained from reliable agricultural databases, research publications, and government reports. Simultaneously, comprehensive information about various ML algorithms was gleaned from academic literature, focusing on their unique requirements, strengths, and potential limitations.

C. *Data Preprocessing*

Upon the successful collection of necessary data, preprocessing was carried out. This stage involved data cleaning to eliminate inaccuracies or inconsistencies and normalization to ensure a standard scale. The pre-processed data were then divided into training and testing sets, thereby setting a robust foundation for the subsequent research stages.

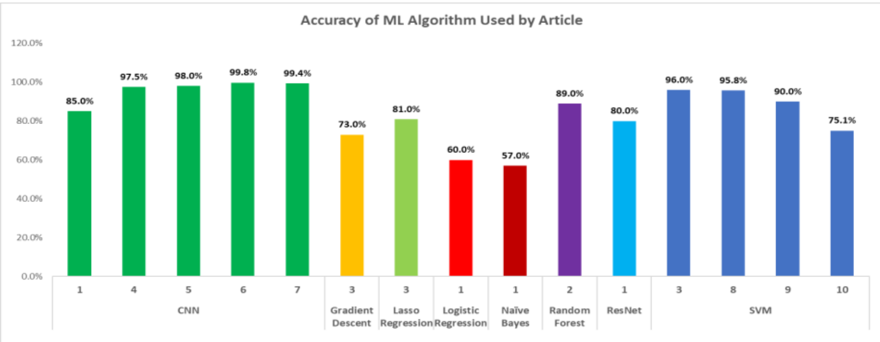

Fig. 1. Accuracy of ML Algorithm Used by Article

D. *Application of the Machine Learning Algorithms*

The subsequent step involved the selection of a diverse set of ML algorithms for evaluation. The chosen algorithms were implemented on the training dataset, followed by performance evaluation and comparison using appropriate accuracy metrics. This process aided in determining the most suitable ML algorithm for application in BambooShoots.AI.

E. *Analysis and Interpretation*

Upon the completion of the ML algorithms evaluation, the research proceeded to the analysis and interpretation phase. The results were thoroughly scrutinized to identify the most effective ML algorithm for BambooShoots.AI, keeping in view the accuracy metrics of the selected algorithm. Limitations discovered during the study were discussed, and potential areas for further improvement were suggested.

III. Results and Discussions

The systematic literature review yielded a wide-ranging perspective on the application of Machine Learning (ML) algorithms in the agricultural Artificial Intelligence (AI) field, specifically concerning bamboo shoots. The ten studies examined in the review each offered unique insights into the problem domain, highlighting both the diversity and potential of ML algorithms applied to this area.

Based on a systematic review of ten research articles, as it is shown in Fig. 1 below, it becomes evident that Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) are the two most frequently utilized machine learning algorithms in the field of plant identification and disease detection. These findings suggest a

high level of confidence among researchers in these particular algorithms' capabilities when dealing with tasks related to agricultural applications.

As demonstrated in the studies under review, CNNs have shown remarkably high performance in tasks related to image recognition. For instance, in the research conducted by Khamparia et al.[9], a Convolutional Neural Network was employed for the prediction and classification of seasonal crop diseases, achieving a high accuracy rate of 97.5%. Similarly, other studies that used CNNs, such as those by Saleem et al.[10] and Karthik et al.[11], reached impressive accuracy rates of 99.81% and 98% respectively. These results highlight the efficacy of CNNs in handling image-based classification problems, making them an appropriate choice for applications like bambooshoot.ai, which presumably deals with image data.

Table 1. Information Of Ten Articles

Author(s)	Year	Title	Journal/Source	DATASET	ML Algorithm Used	Accuracy Metrics
Juyal*, Piyush Kulshrestha, Chitransh Sharma, Sachin Ghanshala, Tejasvi	2020	Common Bamboo Species Identification using Machine Learning and Deep Learning Algorithms	International Journal of Innovative Technology and Exploring Engineering	Forest Research Institute (FRI) Dataset	Naïve Bayes Logistic Regression CNN ResNet	57% 60% 85% 80%
Neelakantan . P,		Analyzing the best machine learning algorithm for plant disease classification	Proceedings, Volume 80, Part 3, 2023, Pages 3668-3671, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.07.358.	220 Images	Random Forest	89%
Jhajharia, Kavita Mathur, Pratistha Jain, Sanchit Nijhawan, Sukriti	2023	Crop Yield Prediction using Machine Learning and Deep Learning Techniques	Procedia Computer Science		Random Forest SVM Gradient Descent Lasso Regression	89% 96% 73% 81%
Khamparia, Aditya Saini, Gurinder Gupta, Deepak Khanna, Ashish Tiwari, Shrasti de Albuquerque, Victor Hugo C.	2020	Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network	Circuits, Systems, and Signal Processing	900 Images	CNN	97.50%
Karthik, R. Hariharan, M. Anand, Sundar Mathikshara, Priyanka Johnson, Annie Menaka, R.	2020	Attention embedded residual CNN for disease detection in tomato leaves	Applied Soft Computing Journal	Plant Village Dataset	CNN	98%
Saleem, Muhammad Hammad Potgieter, Johan Arif, Khalid Mahmood	2020	Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers	Plants	Plant Village Dataset	CNN	99.81%
Mohanty, Sharada P. Hughes, David P. Salathé, Marcel	2016	Using deep learning for image-based plant disease detection	Frontiers in Plant Science	Public Image Dataset 54306 Images	CNN	99.35%
Abdu, Aliyu M. Mokji, Musa M. Sheikh, Usman U.	2020	Machine learning for plant disease detection: An investigative comparison between support vector machine	IAES International Journal of Artificial Intelligence	PV Dataset	SVM	95.75%
Daniya, T. Vigneshwari, S.	2019	A review on machine learning techniques for rice plant disease detection in agricultural	International Journal of Advanced Science and Technology	3000 images	SVM	90%
Kumar*, Sunil Kumar, Vivek Sharma, R.K.	2019	Rice Yield Forecasting using Support Vector Machine	International Journal of Recent Technology and Engineering (IJRTE)	Rice Production of India	SVM	75.06%

On the other hand, SVMs are the second most popular machine learning algorithms applied in the reviewed articles. In the research by Abdu et al.[12], SVMs were used in a comparative study for plant disease detection, and an accuracy of 95.75% was achieved. Similarly, the study by Daniya and Vigneshwari [13] obtained a 90% accuracy rate when SVMs were applied for rice plant disease detection. Despite slightly lower accuracies compared to the CNN-based studies, the SVMs' results are still notably high, indicating that SVMs are a valid and effective alternative for classification tasks in agricultural applications.

Fewer articles applied other algorithms such as Naïve Bayes, Logistic Regression, Gradient Descent, Lasso Regression, and ResNet. While these algorithms also showed promise, the consistently higher accuracies achieved using CNNs and SVMs recommend them for further exploration and consideration. For the experimental analysis that was planned and the subsequent application to bambooshoot.ai, both CNN and SVM were considered as potential algorithm candidates. The final selection was ideally based on further experimental validation using the specific dataset intended for the bambooshoot.ai application. Whichever algorithm ultimately provided the highest accuracy in these experiments was then the recommended choice for the application's development.

A. *CNN Model Architecture*

Based on the systematic review conducted, analyzing ten distinct research papers, it was discerned that Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) were the two predominant machine learning algorithms utilized in the studies. Notably, these two algorithms surfaced as the primary candidates for further experimentation, attributable to their consistent performance and high accuracy metrics across various applications.

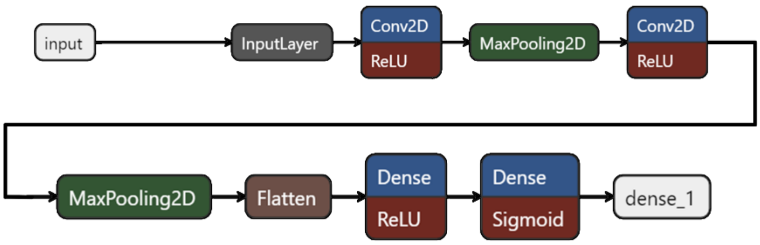

Fig. 2. CNN Model Architecture Diagram

Significantly, CNN was selected as the first algorithm to be trained. As demonstrated in Fig. 2, the intricate architecture of CNN comprises multiple layers that enable it to adeptly detect complex patterns in multidimensional spaces. The figure above showed the CNN model was set up to handle color images of bamboo shoots, each 64x64 pixels in size. As displayed in Fig. 2, the model worked by first applying 32 filters to the image, creating 'feature maps' that emphasized different aspects. This enabled the model to detect basic patterns like edges or textures. To simplify the data without losing important pattern recognition, a 'max pooling' layer was added. This layer also allowed the model to recognize patterns despite minor shifts or rotations. The process was repeated: applying another layer of filters for more complex patterns and a subsequent max pooling layer for simplification. This cycle helped the model learn more intricate details. Once the process was completed, the feature maps were transformed into a line of data using a 'flattening' layer. Traditional neural networks are known as 'fully connected layers' then processed this data, learning to identify combinations of patterns that signified different bamboo shoot conditions.

The final layer in the CNN made a prediction using a function that produced a value between 0 and 1, signifying the likelihood of a specific bamboo shoot condition. The CNN model was trained using the Adam optimization algorithm and binary cross entropy over 25 rounds. This method allowed the model to learn and improve its ability to correctly identify bamboo shoots gradually. In the final test, the model accurately predicted the bamboo shoot conditions based on individual images.

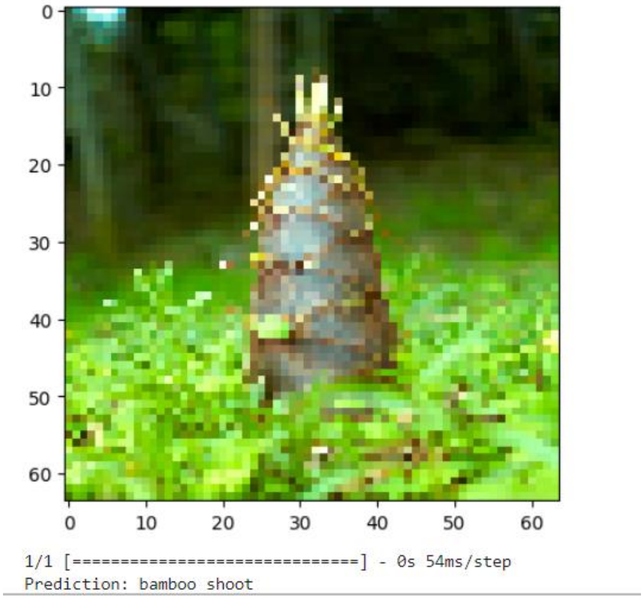

Fig. 3. Sample predicted image using CNN

B. SVM Model Architecture Diagram

The researcher also explored the Support Vector Machine (SVM) as was the second algorithm that had been identified from the systematic review. This algorithm played a significant role in developing the Bambooshoots.Ai.

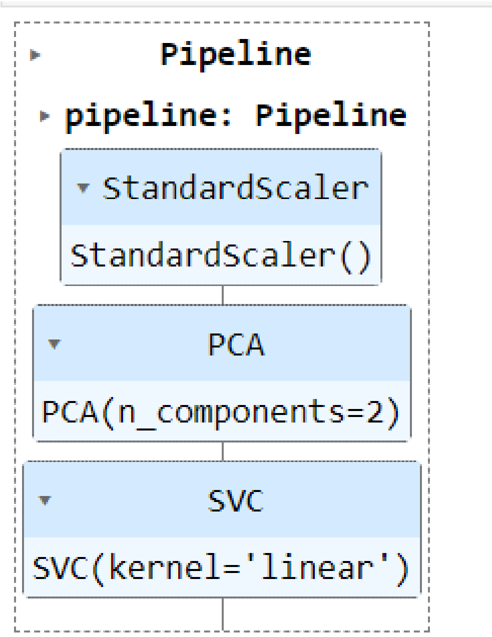

Fig. 4. SVM Model Architecture Diagram

The implementation of SVM started with the introduction of bamboo shoot images, which constituted the primary data for the algorithm. The first stage involved preprocessing, where each image underwent essential transformations, such as resizing, normalization, and grayscale conversion if color information was deemed insignificant.

It's worth emphasizing that the quality of image acquisition and preprocessing directly influenced the accuracy of the subsequent classification process.

Following this, the algorithm embarked on feature extraction, a crucial phase where meaningful attributes or 'features' were derived from the images to represent each bamboo shoot. These features could include texture, shape, color, or any discernible patterns on the bamboo shoots that were potentially linked to their readiness for harvest or the presence of pests/diseases. A carefully designed feature extraction strategy was paramount for the successful training of the SVM model.

After extracting the features, they were fed into the SVM model, which was trained to classify the bamboo shoots into distinct categories such as 'ready for harvest', 'not ready', 'healthy', and 'diseased'. The SVM algorithm aimed to identify an optimal hyperplane in the multidimensional feature space that could effectively separate these categories. SVM accomplished this by maximizing the margin between the support vectors - the data points that were closest to the decision boundary. Importantly, a larger margin was associated with a lower generalization error of the classifier.

The trained SVM model exhibited an accuracy of approximately 57.14%, indicating that more than half of the new bamboo shoot images were correctly classified. While there is room for improvement, this accuracy underscores the potential of the SVM model in image classification tasks within the context of BambooShoots.AI.

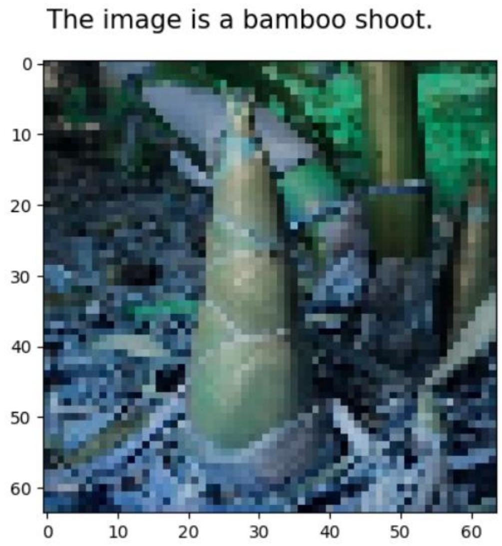

Fig. 5. Sample predicted image using SVM

The trained SVM model exhibited an accuracy of approximately 57.14%, indicating that more than half of the new bamboo shoot images were correctly classified. While there is room for improvement, this accuracy underscores the potential of the SVM model in image classification tasks within the context of BambooShoots.AI.

‘One of the significant benefits of the SVM algorithm was its adaptability. The application of different kernel functions allowed for data transformation into higher-dimensional spaces, thereby managing complex and non-linear classification problems. For instance, if the readiness and health statuses of the bamboo shoots could not be linearly separated in the initial feature space, a suitable kernel function was employed to facilitate separation in a transformed space.

Once the training and validation stages were successfully completed, the model was deployed to predict the readiness for harvest and the health status of new bamboo shoot images. This automation of the system significantly simplified the bamboo shoot farming process, empowering farmers to make knowledgeable decisions, enhancing productivity, and safeguarding crop health.

This research, which assessed SVM as a second algorithm, offered a strong and innovative approach to revolutionizing bamboo shoot farming within the BambooShoots.AI framework.

C. *Test Model Performance through Accuracy Metrics*

This research embarked on an extensive investigation for BambooShoots.AI, assessing two different Machine Learning algorithms for their effectiveness - the Support Vector Machine (SVM) and the Convolutional Neural Network (CNN). The comparison of their accuracies revealed notable differences, which played a significant role in the final decision on the most suitable algorithm for BambooShoots.AI.

```
]: # Evaluate the model
   print("Train Accuracy: ", pipe_model.score(X_train, y_train))
   print("Test Accuracy: ", pipe_model.score(X_test, y_test))

   Train Accuracy:  0.6443298969072165
   Test Accuracy:  0.5714285714285714
```

Fig. 6. Support Vector Machine Accuracy Metrics

The SVM, as the second algorithm assessed, provided an accuracy of approximately 57.14% as shown in Fig. 6, suggesting that it correctly classified a little over half of the new bamboo shoot images. Though a decent starting point, the accuracy indicates considerable room for improvement, especially when compared to the CNN model.

On the other hand, the CNN model, trained over 25 epochs with 8 batches of data per epoch, demonstrated superior performance. There was a consistent increase in training accuracy from an impressive 97.94% in the first epoch as depicted in Fig. 7. Simultaneously, the training loss showed a considerable decrease from 0.0674 in the first epoch to 0.0146 by the completion of the 25th epoch. These results underscore the model's ability to learn effectively from the training data, steadily improving its predictions across successive epochs.

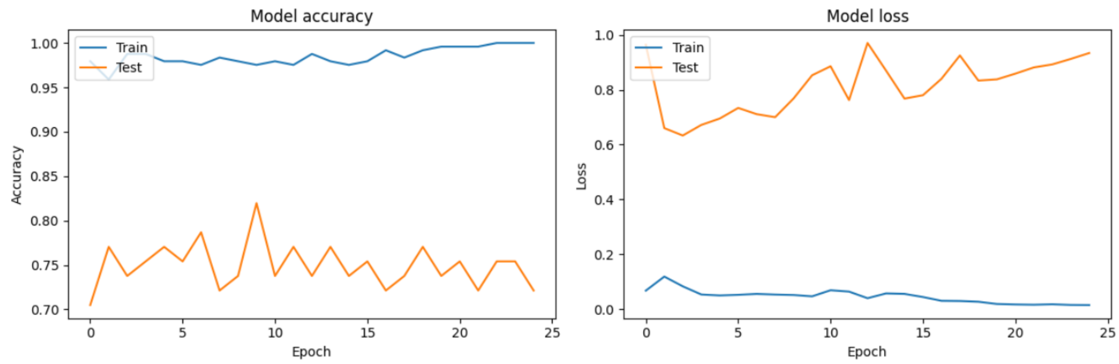

Fig. 5. Convolutional Neural Network Accuracy Metrics

Given the comparison of accuracies, it is evident that the CNN model outperformed the SVM in this specific image classification task for BambooShoots.AI. The significant difference in accuracy levels suggests that the CNN model learned the features of the images more effectively, and hence was better suited to classifying the bamboo shoot images.

CNNs are particularly effective for image classification tasks as they automatically learn hierarchical feature representations. This eliminates the need for manual feature extraction, a process that is necessary in SVM. Given the nature of the task - which involves interpreting complex patterns in images to classify bamboo shoots based on readiness for harvest and the presence of pests/diseases - CNN's ability to handle such intricate image features can be extremely valuable. This, coupled with the significantly higher accuracy of the CNN model, makes it the preferred choice for the BambooShoots.AI application.

IV. Conclusions

In conclusion, the study carried out a thorough investigation into machine learning, investigating the potential algorithms suitable for application in the agricultural AI sector, particularly for bamboo shoots. The systematic review of ten distinct articles offered valuable insights into the use of ML algorithms in the agriculture sector, demonstrating the breadth of application and highlighting the potential of these computational tools.

The review revealed that Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) were the two most frequently applied machine learning algorithms in the field of plant identification and disease detection. Both these algorithms, with their impressive performance metrics in the reviewed studies, warranted further investigation and experimentation in the context of the Bambooshoots.ai application.

Experimentally, it was the CNN algorithm that outshone others, particularly when compared to the SVM. The accuracy of the CNN model consistently increased, reaching a zenith of 97.94% over 25 epochs, a stark contrast to the relatively modest 57.14% accuracy rate of the SVM. This difference indicated the CNN model's superior learning and prediction ability in the classification of bamboo shoot images, a task involving the interpretation of intricate patterns in images.

Hence, based on the evidence from the systematic review and subsequent experimentation, the CNN algorithm emerged as the recommended choice for the development of the Bambooshoots.ai application. Its excellent performance, high accuracy rate, and ability to effectively handle complex image features make it a valuable tool in the AI-based agricultural sector.

References

[1] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” Comput. Electron. Agric., vol. 147, no. February, pp. 70–90, 2018, doi: 10.1016/j.compag.2018.02.016.

[2] N. Zhang, M. Wang, and N. Wang, “Precision agriculture - A worldwide overview,” Comput. Electron. Agric., vol. 36, no. 2–3, pp. 113–132, 2002, doi: 10.1016/S0168-1699(02)00096-0.

[3] P. Rejekiningrum and Y. Apriyana, "Design and implementation of solar pump irrigation systems for the optimization of irrigation and increase of productivity," IOP Conf. Ser. Earth Environ. Sci., vol. 622, no. 1, 2021, doi: 10.1088/1755-1315/622/1/012046.

[4] B. Espejo-Garcia, I. Malounas, E. Vali, and S. Fountas, "Testing the Suitability of Automated Machine Learning for Weeds Identification," Ai, vol. 2, no. 1, pp. 34–47, 2021, doi: 10.3390/ai2010004.

[5] E. J. A. V. Pascual, J. M. J. Plaza, J. L. L. Tesorero, and J. C. de Goma, "Disease detection of Asian rice (Oryza sativa) in the Philippines using image processing," ACM Int. Conf. Proceeding Ser., pp. 131–135, 2019, doi: 10.1145/3366650.3366676.

[6] T. Domingues, T. Brandão, and J. C. Ferreira, "Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey," Agric., vol. 12, no. 9, pp. 1–23, 2022, doi: 10.3390/agriculture12091350.

[7] Neelakantan. P, "Analyzing the best machine learning algorithm for plant disease classification," Mater. Today Proc., no. xxxx, 2022, doi: 10.1016/j.matpr.2021.07.358.

[8] H. Pallathadka, M. Mustafa, D. T. Sanchez, G. Sekhar Sajja, S. Gour, and M. Naved, "IMPACT OF MACHINE learning ON Management, healthcare AND AGRICULTURE," Mater. Today Proc., no. July, 2021, doi: 10.1016/j.matpr.2021.07.042.

[9] A. Khamparia, G. Saini, D. Gupta, A. Khanna, S. Tiwari, and V. H. C. de Albuquerque, "Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network," Circuits, Syst. Signal Process., vol. 39, no. 2, pp. 818–836, 2020, doi: 10.1007/s00034-019-01041-0.

[10] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers," Plants, vol. 9, no. 10, pp. 1–17, 2020, doi: 10.3390/plants9101319.

[11] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, "Attention embedded residual CNN for disease detection in tomato leaves," Appl. Soft Comput. J., vol. 86, p. 105933, 2020, doi: 10.1016/j.asoc.2019.105933.

[12] A. M. Abdu, M. M. Mokji, and U. U. Sheikh, "Machine learning for plant disease detection: An investigative comparison between support vector machine and deep learning," IAES Int. J. Artif. Intell., vol. 9, no. 4, pp. 670–683, 2020, doi: 10.11591/ijai.v9.i4.pp670-683.

[13] T. Daniya and S. Vigneshwari, "A review on machine learning techniques for rice plant disease detection in agricultural research," Int. J. Adv. Sci. Technol., vol. 28, no. 13, pp. 49–62, 2019.

[14] K. Jhajharia, P. Mathur, S. Jain, and S. Nijhawan, "Crop Yield Prediction using Machine Learning and Deep Learning Techniques," Procedia Comput. Sci., vol. 218, pp. 406–417, 2023, doi: 10.1016/j.procs.2023.01.023.

[15] S. Kumar*, V. Kumar, and R. K. Sharma, "Rice Yield Forecasting using Support Vector Machine," Int. J. Recent Technol. Eng., vol. 8, no. 4, pp. 2588–2593, 2019, doi: 10.35940/ijrte.d7236.118419.

[16] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Front. Plant Sci., vol. 7, no. September, pp. 1–10, 2016, doi: 10.3389/fpls.2016.01419.

[17] P. Juyal*, C. Kulshrestha, S. Sharma, and T. Ghanshala, "Common Bamboo Species Identification using Machine Learning and Deep Learning Algorithms," Int. J. Innov. Technol. Explor. Eng., vol. 9, no. 4, pp. 3012–3017, 2020, doi: 10.35940/ijitee.d1609.029420.