Groundnut Disease Detection with Efficient Deep Learning and Model Compression

Sandra Jeslena Wesley, Getzi Jeba Lellipushpam Paulraj,
Immanuel Johnraja Jebadurai

Karunya Institute of Technology and Sciences, sandrajeslena@gmail.com

Abstract.

Groundnut diseases have notably influenced agricultural productivity over the years. This work, therefore, has put under experimentation the applicability of DenseNet-state-of-the-art convolutional neural network architecture in groundnut disease prediction. The research work has been extended to pruning techniques to enable real-time inference with reduced latency and resource consumption. The model is trained using a dataset of 10,361 annotated images of groundnut leaves, categorized into six classes. The results prove the effectiveness of the proposed pruned methodology in terms of accuracy, inference speed, floating operations, memory efficiency, and the introduction of scalable solutions in agriculture. It is also observed that the global pruning shows better performance results along with an improved accuracy of 2%.

Keywords. DenseNet Architecture, Precision Agriculture, Auto-compression Pruning Model, Convolutional Neural Networks (CNNs), Pruned Neural Networks, and Real-Time Inference.

1. INTRODUCTION

Groundnut is one of the significant crops in the world, being a major source of protein and oil. Groundnut diseases are responsible for substantial economic losses worldwide. They decrease the yield and quality, affecting both farmers and industries that depend on groundnuts as raw materials. Diseases like leaf spot, rust, and stem rot are serious issues and threats to its production. These diseases, due to their nature, are usually unnoticed until the plant has been badly affected by a disease, which in turn reduces productivity and economic returns. Early detection and treatment can avoid these losses, to ensure a stable income for farmers and a reliable supply for the market.

This paper deals with [1] Deep learning with CNNs, improving in value by transfer learning, offering high accuracy and efficiency in plant disease detection, and imposing pre-trained models for small, domain-specific datasets. Among the pre-trained models, DenseNet has attracted wide attention with its efficacy in feature extraction and reuse. Unlike conventional convolutional networks, DenseNet connects each layer with every other layer, thereby enhancing the flow of gradients with reduced vanishing gradients, and making it much more suitable in cases that require finding very complicated patterns. While DenseNet offers high accuracy, its computational complexity can make slow real-time applications.

This paper proposes a lightweight DenseNet model that performs early detection of groundnut diseases. Performance analysis shows that the proposed lightweight model outperforms the traditional DenseNet model in terms of accuracy, number of learning parameters, and Floating Points Operations Per Second (FLOPS).

2. RELATED WORKS

Detecting diseases and increasing crop yields are only two of the issues in agriculture that deep learning has made possible [1]. Traditional methods like K-means clustering and Support Vector Machines (SVMs) were used in plant disease detection by analyzing the color, texture, and shape features of leaf images. These methods, however, endure from the accuracy and reliability of the features [2]. In comparison, deep learning methodologies, specifically Convolutional Neural Networks (CNNs), have fared much better with feature learning automatically applicable to a problem in question [3,4]. Methods including AlexNet, GoogLeNet, and DenseNet have been extensively adopted for such an intention. VGG16 and ResNet, for example, have supported accuracy gain over small and unbalanced datasets, and visualization tools including heatmaps and saliency maps have aided in delivering explicit information regarding model operations.

DenseNet is a deep model that maximizes information flow in a network via connectivity in layers Thus, feature use is optimized, vanishing gradients are eliminated, and overall requirements for parameters are reduced. Variants of DenseNet [5], such as integration with residual connectivity and new types of multiplication, have continued to make its feature extraction even more powerful and with reduced computational requirements. Thus, DenseNet is an ideal model for implying high accuracy with high efficiency [6,7]. The quality of datasets is critical for disease detection in plants, as well. Public datasets such as Plant Village, and in-house datasets acquired through hyperspectral imaging [8], and even through a manual process, are utilized regularly. The use of deep neural networks in mobiles or constrained devices is challenging in terms of model size. Model pruning, model quantization, and knowledge distillation can reduce model size with a minor loss in accuracy. Techniques such as genetic algorithms can minimize noise, generating cleaner images for a boost in accuracy in detection. Optimizers such as Rectified Adam (RAdam) [9] and focal loss, a personalized loss function, stabilize training and accuracy even better, even in unbalanced and noisy datasets.

3. PROPOSED METHODOLOGY

The workflow includes model architecture design, optimization strategies, and real-time deployment considerations. To kickstart the workflow of the architecture, a specifically crafted dataset was used for the implementation. The raw dataset initially consisted of 3,059 images, which were further processed and augmented to enhance its robustness for deep learning applications. Augmentation techniques such as cropping, rotation, flipping, brightness adjustments, and contrast enhancement were applied, leading to an expanded dataset of 10,361 labeled images. To maintain consistency and optimize computational efficiency, all images were resized to 256 × 256 pixels.

The proposed method has the following two modules:

- DenseNet Backbone: The idea is based on the efficiency of DenseNet for feature extraction and also connects layers to capture complex patterns of groundnut leaves.
- Auto-Compression Pruning: Introduce the pruning mechanism that gets rid of extra parameters and brings in lightweight models without compromising accuracy.

3.1 ARCHITECTURE OF DENSENET

DenseNet consists of multiple dense blocks, where each layer receives input from all preceding layers and passes its output to all subsequent layers. This results in L(L+1)/2 direct connections for L layers, significantly enhancing information flow. The architecture comprises:

Dense Blocks: Each block consists of multiple convolutional layers with batch normalization and ReLU activation. Feature maps from all preceding layers are concatenated, promoting information retention, Transition Layers: Used between dense blocks to reduce the number of feature maps and downsample spatial dimensions through batch normalization, 1x1 convolutions, and average pooling, Growth Rate (k): Defines the number of features maps each layer generates, influencing computational complexity and network capacity.

3.2 AUTO-COMPRESSION PRUNING

Auto-compression Pruning Model:

The architecture of the proposed system for groundnut disease detection combines the DenseNet framework with auto-compression pruning techniques to create a robust, efficient, and scalable model.

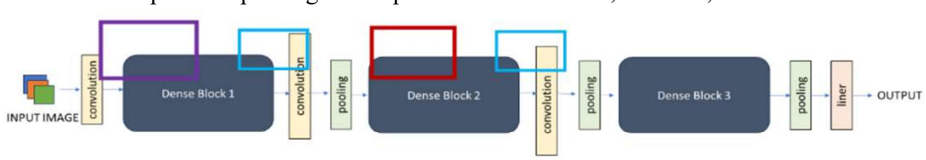


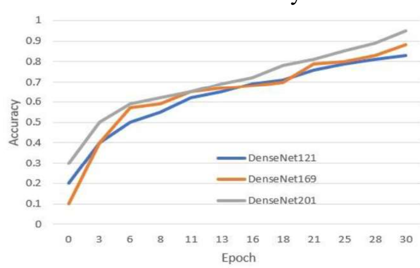
Figure 14 DenseNet Auto-Compression Pruning

Pruning Occurs:

Inside Each Dense Block:
Pruning removes less important feature maps and connectivity in convolutional layers for model and computation size reduction. In the present image(fig.1), pruning will occur between feature maps in Dense Block 1, Dense Block 2, and Dense Block 3.
Between Dense Blocks (Transition Layers):
The Pooling + Convolution layers act to bridge between Dense Blocks. In such a scenario, pruning reduces unnecessary channels in anticipation of output delivery to the successive block.

4. IMPLEMENTATION

A comparative study of DenseNet-121, DenseNet-169, and DenseNet-201 was done to obtain the best model for detecting diseases in groundnut plants. The performance is compared considering test dataset classification accuracy for all three architectures on the same protocol of training and validation. From the graph (fig.2) in support, it is understood that DenseNet-121 has an accuracy of 91%, DenseNet-169 has an accuracy of 95% and DenseNet-201 has an accuracy of 97%.



Therefore, the incremental gain from DenseNet-121 up to DenseNet-201 proves the worth of deep architectures in complex disease feature extraction. The reason for this superiority could be easily understood by having more layers within the DenseNet-201. That means DenseNet-201 can learn more elaborated patterns, making it the fittest candidate for our task.

Figure 2 Comparative study graph

5. RESULTS AND DISCUSSION

Impact of Pruning Techniques on Model Performance
Pruning is a crucial model compression technique that reduces redundancy in deep learning models while maintaining accuracy and efficiency. This study evaluates the impact of various pruning techniques—Structured Pruning [11], Unstructured Pruning [11], Layered Pruning [12], Filter Pruning [13], and Global Pruning—on model performance, computational efficiency, and resource consumption.

Metric	Before Pruning	Structured Pruning	Unstructured Pruning	Layered Pruning	Filter Pruning	Global Pruning
Accuracy	72.00%-91.33%	71.83%	82.67%	74.00%	71.67%	94.17%
Model Size	77.57 MB	67.83 MB	67.16 MB	75.16 MB	75.83 MB	68.16 MB
FLOPs	140.49 GFLOPs	133.49 GFLOPs	140.49 GFLOPs	140.49 GFLOPs	140.49 GFLOPs	140.49 GFLOPs
Parameters	20.01M	18.01M	20.01M	20.01M	20.01M	18.01M
Inference Time (sec/image)	0.178054-0.229022	0.163522	0.177324	0.193090	0.171260	0.161881

DISCUSSION
Comparison of Model Before and After Pruning:
Among all pruning techniques, global pruning emerged as the most effective method. While others such as structured, unstructured, layered, and filter pruning led to accuracy drops, global pruning uniquely improved accuracy from 91.33% to 94.17%, and further to 96.83% in the final pruned model. The original model size of 77.57 MB was significantly reduced through pruning techniques, with Unstructured Pruning achieving 67.16 MB and Global Pruning reducing it to 68.16 MB, making it the most efficient. The initial inference time ranged from 0.178 to 0.229 seconds per image, with Global Pruning achieving the fastest inference time at 0.161 seconds per image, followed by Structured Pruning at 0.163 seconds per image. Global Pruning also reduced memory usage to 95.13 MB, making it the best choice for low-power devices. Additionally, Structured Pruning reduced FLOPs to 133.49 GFLOPs, while other pruning techniques retained the original 140.49 GFLOPs. The mention of pruning techniques for model compression is valuable; however, explicitly highlighting their impact on model size reduction and inference speed would strengthen the contribution, demonstrating how pruning enhances model efficiency for real-time applications. After pruning, parameters were reduced to 12,741,857, and the total compute cost dropped to 96.09 GMac.

1. Global pruning was the most effective technique, maintaining model accuracy with minimal performance loss, while other methods (structured, unstructured, layered, and filter pruning) led to significant accuracy drops.
2. The final pruned model reaches a higher accuracy (96.83%) than the unpruned version (94.00%), demonstrating the potential of efficient pruning strategies to enhance model performance while optimizing computational efficiency.

6. CONCLUSION

The implementation of DenseNet with global pruning and model compression put forward a transformative solution for groundnut disease detection, addressing computational inefficiency and deployment challenges on low-power devices. Global pruning effectively removes redundant parameters while preserving high accuracy, unlike structured and unstructured pruning, which often degrade performance. By integrating DenseNet with pruning, the model achieves reduced memory and computational complexity, enabling efficient inference on edge devices with minimal latency. This enhances real-time disease detection, making AI-driven solutions accessible to farmers in rural areas. This approach not only improves crop monitoring and early disease intervention but also promotes sustainable farming by reducing pesticide use. Future work could focus on adaptive pruning strategies to optimize performance based on dataset and hardware constraints, further advancing scalable and efficient agricultural disease detection systems.

7. REFERENCES

[1] Lili Li, Shujuan Zhang, Bin Wang, "Plant disease detection and classification by deep learning—a review," in IEEE Access, vol. 9, pp. 56683-56698, 2021.

[2] Bolin Chen, Tiesong Zhao, Jiahui Liu, Liqun Lin, "Multipath feature recalibration DenseNet for image classification," in International Journal of Machine Learning and Cybernetics, vol. 12, pp. 651-660, 2021.

[3] Sachin D. Khirade, Amit B. Patil, "Plant disease detection using image processing," in 2015 International Conference on Computing Communication Control and Automation, pp. 768-771, 2015.

[4] Konstantinos P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," in Computers and Electronics in Agriculture, vol. 145, pp. 311-318, 2018.

[5] Lifeng Yin, Pujiang Hong, Guanghai Zheng, Huayue Chen, Wu Deng, "A novel image recognition method based on DenseNet and DPRN," in Applied Sciences, vol. 12, no. 9, article 4232, 2022.

[6] Tavishee Chauhan, Hemant Palivela, Sarveshmani Tiwari, "Optimization and fine-tuning of DenseNet model for classification of COVID-19 cases in medical imaging," in International Journal of Information Management Data Insights, vol. 1, no. 2, article 100020, 2021.

[7] Najmul Hasan, Yukun Bao, Ashadullah Shawon, Yanmei Huang, "DenseNet convolutional neural networks application for predicting COVID-19 using CT images," in SN Computer Science, vol. 2, no. 5, article 389, 2021.

[8] Guandong Li, Chunju Zhang, Runmin Lei, Xueying Zhang, Zhourun Ye, Xiaoli Li, "Hyperspectral remote sensing image classification using three-dimensional-squeeze-and-excitation-DenseNet (3D-SE-DenseNet)," in Remote Sensing Letters, vol. 11, no. 2, pp. 195-203, 2020.

[9] Zheng Wan, Zhang Yuxiang, Xuhui Gong, Boyang Yu, "DenseNet model with RAdam optimization algorithm for cancer image classification," in 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 771-775, 2021.

[10] Y. He and L. Xiao, "Structured Pruning for Deep Convolutional Neural Networks: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 5, pp. 2900-2919, May 2024, doi: 10.1109/TPAMI.2023.3334614.

[11] A. Jordao, M. Lie and W. R. Schwartz, "Discriminative Layer Pruning for Convolutional Neural Networks," in IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 4, pp. 828-837, May 2020, doi: 10.1109/JSTSP.2020.2975987.

[12] J. -H. Luo, H. Zhang, H. -Y. Zhou, C. -W. Xie, J. Wu and W. Lin, "ThiNet: Pruning CNN Filters for a Thinner Net," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 10, pp. 2525-2538, 1 Oct. 2019, doi: 10.1109/TPAMI.2018.2858232.

Biographies

Sandra Jeslena Wesley is currently pursuing her B.Tech degree in Computer Science and Engineering at Karunya Institute of Technology and Sciences, Coimbatore, India, as part of the 2025 batch. She has practical knowledge in Android programming, deep learning, and NLP, with experience working on applications such as a job recommendation system and an event management platform. She has gained industry exposure through internships at CISCO and IIT, where she worked on research and technology-driven projects. Her areas of interest include AI model development using DenseNet, pruning, and quantization to enhance model efficiency. She aims to apply her skills to develop scalable and practical software solutions.

Getzi Jeba Lellipushpam Paulraj received her B.E. degree in Electronics and Communication Engineering from M.S. University, Tirunelveli, India in the year 2004. She received her M.Tech. Degree in Network and Internet Engineering from Karunya University, Coimbatore, India in the year 2009. She received her Ph.D. from Karunya Institute of Technology and Sciences, Coimbatore, India in the year 2018. Her areas of interest include IoT, Fog Computing, and Data Analytics.

Immanuel Johnraja Jebadurai received his B.E. degree in Computer Science and Engineering from M.S. University, Tirunelveli, India. in the year 2003. He received his M.E. degree in Computer Science and Engineering from Anna University, Chennai India in the year 2005. He received his Ph.D. in Computer Science and Engineering from Karunya Institute of Technology and Sciences, Coimbatore, India in the year 2017. His areas of interest include Network Security, Vehicular Ad-Hoc Networking, and IoT.