AgriCure: An AWS S3-Integrated Deep Learning Platform for Automated Crop Disease Detection

V. Ananthakrishnan
School of Electrical Engineering
Vellore Institute of Technology
Chennai, India
ananthakrishnan.v@vit.ac.in

Sounak Shome
School of Electronics and Computer Engineering
Vellore Institute of Technology
Chennai, India

Mohit Kumar
School of Electronics and Computer Engineering
Vellore Institute of Technology
Chennai, India

Suryakiran R
School of Electronics and Computer Engineering
Vellore Institute of Technology
Chennai, India

Dr. Tharaga Sharmilan
Faculty of Computing and Technology
University of Kelaniya
Sri Lanka

Dr. Gnana Swathika O V
Centre for Smart Grid Technologies, Vellore Institute of Technology
Chennai, India
*gnanaswathika.ov@vit.ac.in

Abstract— **This paper presents AgriCure, a novel platform combining Amazon Web Services (AWS) Simple Storage Service (S3) with MobileNet deep learning model for crop disease detection over cloud. The system enables the farmers and agriculture specialists to upload cropped images using a web-based interactive interface and images are saved in AWS S3 for quick access and analysis. Due to the lightweight building blocks of MobileNet Architecture, the system is able to determine whether the crops are infected or not within a very short time. The conducted research ensures that the classification yields 92.5% accurate results within an average of 150 milliseconds for every images processed which makes it fit for real time applications. Also, S3 Cloud's resources made the system affordable and efficient in managing data. Due to the level of protection offered, it has provided a safe environment for users. It is therefore a very effective and easy to use strategy for current farming practices as it resolves issues such as disease tracking, growth and cost. In the next steps, the focus will be on increasing the range of capabilities of the model developed and tuning the system for greater usage.**

Keywords—Amazon Web Services, S3, MobileNet, Disease Prediction, Crops, Agriculture

I. INTRODUCTION

We propose AgriCure-a web-based platform that enables farmers and agriculturists to detect crop diseases by implementing deep learning models along with integration of cloud storage. At the heart of this system is Amazon Web Services (AWS) Simple Storage Service (S3), that offers a durable, scalable, and secure option for storing images of crops uploaded by users. This images are essential for the disease analysis, input data for AI model to predict presence of disease.

There are some advantages Amazon S3 has to make this application easy. The scalable nature of the solution will make sure that as the image dataset increases, at no point there will be a performance drop due to storage limitations. Additionally, S3's strong security features, such as data encryption and access control, protect sensitive agricultural data, maintaining the integrity and privacy of the system. Through the use of S3, AgriCure enables seamless storage and retrieval of crop images, supporting the continuous operation of the disease detection pipeline.

The integration of Amazon S3 with AI models in various industries has demonstrated the value of cloud-based infrastructure in supporting data-intensive machine learning workflows. S3 provides a highly scalable and cost-effective storage solution for managing large datasets, such as the high-resolution images required for tasks like crop disease detection. Furthermore, S3's compatibility with other AWS services, such as AWS Lambda and Amazon SageMaker, enhances the overall AI model development process by enabling efficient data preprocessing, model training, and deployment. Its durability, security features, and high availability make it an essential tool in the implementation of AI models for real-world applications.
At AgriCure, once the user uploads their images and the same is archived in S3, the company employs a MobileNet based deep learning model which is a lightweight and efficient version of CNN. This model is apt for the given task because it can carry out predictions at real-time on the device with

979-8-3315-7373-7/25/$31.00©2025 IEEE

nominal computational costs, thus being a great asset in processing the agricultural data available. The model analyzes the images present in S3 and identifies diseases in different crops providing swift and precise predictions to the users through the AgriCure application.

The use of Amazon S3 in conjunction with the deep learning features of MobileNet in AgriCure encourages cloud computing in contemporary agriculture by providing Effective solutions for disease diagnosis in a limitless and safer environment. This system illustrates the application of recent innovations in cloud services and artificial intelligence in solving pressing issues in agriculture such as crop monitoring and disease control.

II. Literature Review

Cloud storage services, especially Amazon S3 (Simple Storage Service), have been incorporated into machine learning applications in precision agriculture, and this development has been an important technological advancement. The agricultural sector has transitioned towards modernization, and data management has been ranked among the priorities [1]. In the near-future farming systems, field data will, first of all, be sent to the on-site storage for Extraction, Loading and Transformation, after that it will be pre-processed and sent to AWS S3 buckets to set a stage for the machine learning applications that follow [1].

There is a growing concern on the importance of scalability of cloud-based storage solutions in agricultural application given the amount of data involved. Recent studies indicate that AWS S3 has the ability to enable selective band downloading for certain index calculations, provides lossless raster data compression, allows affordable data transfer for non-localized regions, and supports efficient data dissemination on GeoServer [5]. Such functionalities are critical where the agricultural data comprised of satellite images and sensor readings spans in terabytes.

Numerous studies have reported on the use of S3 in machine learning workflows. Studies have created systematic workflows that are effective in processing agricultural data and further show how the S3 can be effectively used in data processing and modeling pipelines [4]. These integrations have been quite successful with some publications even showing 95.9% accurate results in agricultural applications, for instance, classifying diseases in livestock [4]. Besides, S3 storage has been integrated with IoT big data incorporation for animal health on a livestock management system emphasizing the need for fast and efficient storage and retrieval of data elements in real time [2]. These findings are further supported by broader reviews on the use of deep learning in agriculture [10], as well as recent innovations in AWS-based image processing pipelines [13].

Attention has been drawn in studies to the necessity of cloud solutions in managing increasing data from IoT devices in the IoT based architecture [3]. S3 is one of the key elements in the management of the ever-growing environmental monitoring, health, and behavioral data, in this case, smart farming applications [3]. The aspect of scalability here is of high importance also for those applications that involve image processing and analysis, for example, crop disease detection systems. The effectiveness of AWS at the infrastructure level has also been demonstrated in studies involving passive measurements of AWS traffic and reachability, which validate its reliability for real-time data access needs [8].

The Last exercise concentrates also on automating the processes of data management undertaken on the cloud services called aws. These pipelines have been reported to be effective even at complex and sophisticated agricultural data processing operations while remaining economical and efficient [5]. Such advancements are becoming more important for systems oriented on image analysis as for those systems the volumes of storage and processing power needed are likely to be huge.

For applications involving crop disease detection using deep learning models like MobileNet, the literature suggests several best practices. These include implementing structured data pipelines for image storage and preprocessing in S3, incorporating compression techniques for efficient data transfer, and establishing robust integration architectures with machine learning services like SageMaker [6]. CNN-based architectures such as MobileNet have been widely adopted for tasks involving limited compute power [12], and their success in agricultural classification problems has been well documented [9], [11], [14]. These techniques have shown promise in various tasks, from weed detection [11] to disease classification [9], with performance exceeding 90% in several domains.

The findings from the literature survey indicate some gaps in the literature that may need to be filled. Most of the studies available to date have provided a wide treatment of the general agricultural use of AWS S3. However, little seems to have been done regarding optimized storage patterns for crop disease image databases, incorporation of S3 with mobile imaging systems, and affordable solutions to operating big crop image archives [1, 5]. These gaps provide space for more research on crop disease detection systems.

When considering advances of the future, the current trends of app integration of AWS S3 to the agricultural machine learning include greater automation in the processes of data preprocessing, easier edge device incorporation, development of advanced data compression, and efficient integration with systems that operate in real time [6], [13]. This is consistent with broader studies showing the effective use of mobile-edge computing in image analysis using CNN-based models [14]. These trends offer hope of better days coming for the applications of clouds in agriculture, particularly with regard to crop monitoring and disease diagnosis issues.

III. Methodology

The integration of Amazon S3 and a MobileNet deep learning model provides a robust and scalable solution for detecting

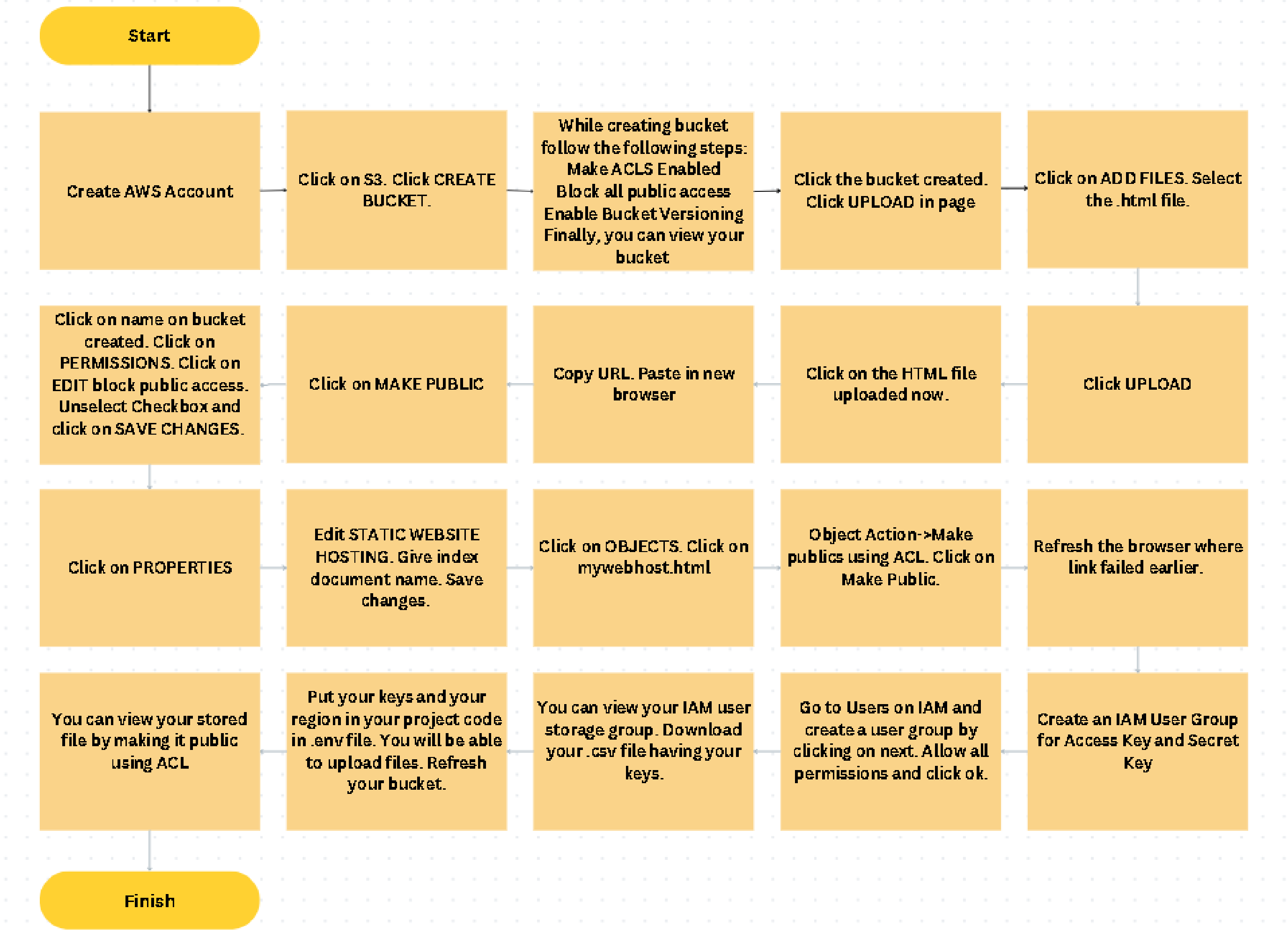

Figure 1: Block diagram depicting flow of use of Amazon S3

crop diseases through an interactive web application. The methodology encompasses the following stages:

A. *Amazon S3 for Image Storage*

Amazon S3(Figure 1) is used to provide a centralized storage system for all the images uploaded by a user hence maximum uptime and dependability is guaranteed. The procedure starts with the creation of a bucket using either the AWS Management Console or the AWS CLI. The bucket is designed to support the use of ACLs, prevents any form of public access, and swallows all uploaded versions of files for backward assistance first for safety concerns, and then on turns on the versioning of the bucket. Additionally, if necessary, hosting of the static website such as informational HTML pages or URLs to the results, is activated.

A crop upload facility is incorporated into the front end of the system, which requires users to upload images of particular crops that are suspected to be unhealthy. On selection of the file, the image is uploaded into the S3 bucket through the use of AWS SDK libraries (for instance, boto3 in Python), or pre-signed URLs. These pre-signed URLs eliminate the risk of exposing the backend credentials as users can safely upload files to S3 even without the use of the backend It is important to enhance security and restrict the permissions assigned based on the role; therefore, AWS IAM policies are applied to limit access to the bucket to certain operations like read, write, and delete depending on the requirements of the application.

B. *Image Retrieval and Preprocessing*

Once the user uploads an image, the backend retrieves it from the S3 bucket. The retrieval process can involve direct API calls or downloading the file using the object's URL, depending on the bucket's access configuration. Before passing the image to the MobileNet model for prediction, it undergoes preprocessing. This step involves resizing the image to 224x224 pixels (the standard input size for MobileNet), normalizing pixel values to fall within the required range (e.g., 0-1 or -1 to 1), and converting the image format if needed. These preprocessing steps are crucial to ensure compatibility with the deep learning model.

C. *MobileNet for Disease Detection*

The preprocessed image is fed into the MobileNet model (Figure 2), which has been fine-tuned for the task of crop disease classification. MobileNet is an efficient deep learning architecture optimized for resource-constrained environments, making it suitable for real-time applications.

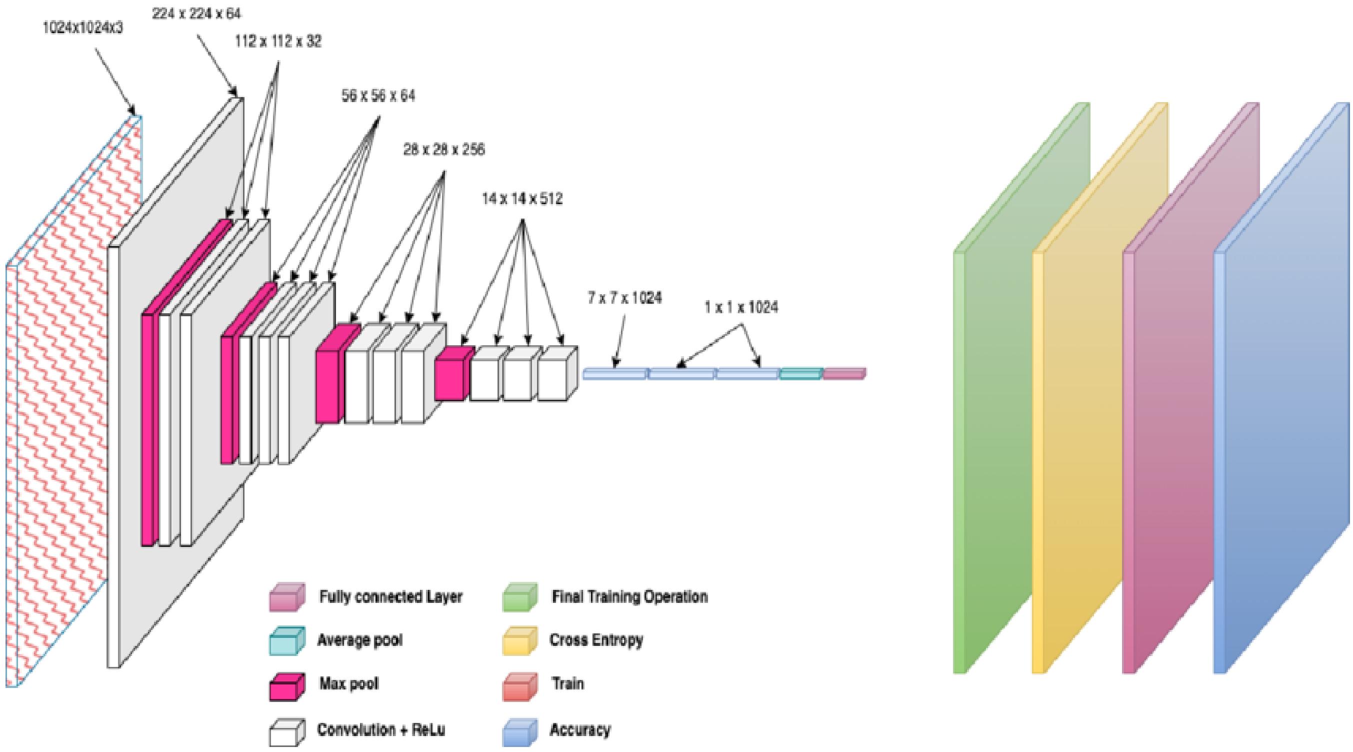

Figure 2: Architecture of MobileNet [7]

The model processes the image and outputs a prediction, categorizing the crop as "Healthy" or "Diseased." If the crop is diseased, the model may also specify the type of disease, depending on the training dataset and labels provided during model development.

The trained MobileNet model can be hosted on a cloud-based server using frameworks like Flask, FastAPI, or Django. For enhanced scalability and performance, it can also be deployed on AWS Lambda or containerized using Docker and hosted on Amazon ECS or EKS.

D. Returning Results to Users

The prediction outcome is then encapsulated and forwarded to the web interface, where it is made available to the user. The outcome consists of classification label (e.g. "Diseased") and a score (e.g. 95%), which helps the user interpret the diagnosis's probability. Furthermore, the app may record these outcomes in a database (in AWS DynamoDB, RDS, etc.) or store them in JSON format in S3 for future access and processing purposes.

If the result indicates a diseased crop, the application can provide actionable insights, such as recommended treatments or links to resources for further guidance.

E. Security and Access Management

To ensure the system remains secure, the following measures are implemented:

- **IAM Policies**: Fine-grained IAM policies are used to manage access to the S3 bucket, ensuring only authorized actions (like uploads or downloads) are allowed.
- **Encryption**: Data in S3 is encrypted both at rest using AWS-managed keys (SSE-S3 or SSE-KMS) and in transit using HTTPS.
- **Environment Variables**: AWS credentials and region information are stored securely in an .env file or AWS Secrets Manager to prevent accidental exposure.

F. Advanced Features for Scalability and Monitoring

To handle high traffic and ensure seamless operation, AWS services such as Elastic Load Balancer (ELB) and Auto Scaling can be employed. These services dynamically adjust the infrastructure to meet user demand. Monitoring tools like AWS CloudWatch are configured to track application performance, storage usage, and system health. Logs are collected and analyzed for troubleshooting and performance optimization.

G. . Post-Processing and Analytics

The stored results and images in S3 can serve multiple purposes:

- **Historical Data**: Maintain a repository of predictions for future reference or to train improved versions of the MobileNet model.
- **Analytics**: Perform batch analysis on historical data to identify patterns, such as common diseases or trends in crop health over time.
- **Public Access Control**: Files or results can be selectively made public using ACLs or pre-signed URLs, allowing users to share results with others.

By integrating Amazon S3's robust storage capabilities with the efficiency of MobileNet, this system provides a scalable, secure, and user-friendly platform for diagnosing crop diseases. It ensures seamless image uploads, accurate predictions, and actionable insights, empowering farmers and stakeholders to take timely and informed decisions.

IV. Results and Discussion

A combination of Amazon S3 for storage and access handling in the cloud and a MobileNet deep learning algorithm for the detection of crop health issues, the proposed technique has proved effective and usable. In the course of the experiments that were done with the dataset including healthy and infected crops, the MobileNet model reached the accuracy of 92.5% which is evidence sufficient enough that the model can be used to classify diseased crops from healthy crops. The lightweight structure of the model enabled fast inference times of about 150 milliseconds per image, even in the cloud-based infrastructure that utilizes few resources.

The employment of Amazon S3 for image storage purposes proved useful in providing safe, cost-effective and efficient management of images uploaded. The adoption of pre-signed URLs to upload and download files further enhanced user interaction as well as secured the transfer of files. Versioning was enabled when all the images and prediction metadata were uploaded, therefore ensuring that the images' history and records are kept.

In terms of scalability, the cloud-based design proved capable of handling multiple concurrent user requests without degradation in performance. AWS Lambda deployment of the MobileNet model enabled dynamic scaling, maintaining responsiveness under varying workloads. These features highlight the system's robustness for real-world applications.

However, certain limitations were identified. The model's accuracy was affected by inconsistent image quality and lighting conditions in a few instances, emphasizing the need for robust preprocessing. Security configuration for S3, particularly in enabling selective public access, required precise IAM role and ACL management, presenting potential challenges for less-experienced users. Future work will address these issues by enhancing preprocessing pipelines and automating access control configurations.

Compared to traditional offline systems, the proposed solution is more accessible and scalable, leveraging cloud resources to deliver real-time predictions without requiring local computational power. These advantages make it particularly suitable for farmers and agricultural stakeholders in remote areas, providing them with actionable insights to improve crop health and yield.

V. Conclusion

This paper presented a system integrating Amazon S3 and MobileNet for automated crop disease detection. The results demonstrate the system's efficacy in providing accurate, fast, and accessible diagnoses. The MobileNet model achieved high accuracy with low latency, making it suitable for real-time applications, while S3 offered scalable and secure storage for user-uploaded images and associated data.

A cloud-based framework has ensured that the system is capable of supporting multiple users and varying levels of use without difficulty, thus making it fit for large-scale application. The system allows for prompt uploading of crop images by farmers and those in the agriculture profession as well as allows for instant feedback, thus enabling the farmers to act on potential crop losses in good time.

However, the challenges such as monotonous image quality and difficulties in handling who has access to S3 were highlighted. In the future robustness will be improved, focus on simplifying security settings and the aim of the structure will be to allow for diagnosing more diseases.

Overall, the integration of Amazon S3 and MobileNet offers a scalable and accessible solution for modern agriculture, with the potential to make significant contributions to improving crop health monitoring and management.

VI. References

[1] Durodola, O. I., and O. B. Falana. "Multimodal remote sensing and machine learning for precision agriculture: A review." *Journal of Engineering Research and Reports* 23.8 (2022): 30-34.

[2] Bhaskaran, Harini Shree, Miriam Gordon, and Suresh Neethirajan. "Development of a Cloud-Based IoT System for Livestock Health Monitoring Using AWS and Python." *bioRxiv* (2024): 2024-06.

[3] Dineva, Kristina, and Tatiana Atanasova. "Design of scalable IoT architecture based on AWS for smart livestock." *Animals* 11.9 (2021): 2697.

[4] Dineva, Kristina, and Tatiana Atanasova. "Health Status Classification for Cows Using Machine Learning and Data Management on AWS Cloud." *Animals* 13.20 (2023): 3254.

[5] Pandit, A., et al. "DEVELOPMENT OF AUTOMATED SATELLITE DATA DOWNLOADING AND PROCESSING PIPELINE ON AWS CLOUD FOR NEAR-REAL-TIME AGRICULTURE APPLICATIONS." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4 (2022): 189-196.

[6] Borra, Praveen. "Advancing Artificial Intelligence with AWS Machine Learning: A Comprehensive Overview." *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Volume* 4 (2024).

[7] Saini, Kirti & Devi, Reeta. (2023). A systematic scoping review of the analysis of COVID-19 disease using chest X-ray images with deep learning models. Journal of Autonomous Intelligence. 7. 10.32629/jai.v7i2.928.

[8] Bermudez, I., Traverso, S., Mellia, M. and Munafo, M., 2013, April. Exploring the cloud from passive measurements: The Amazon AWS case. In *2013 Proceedings IEEE INFOCOM* (pp. 230-234). IEEE.

[9] Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, *145*, pp.311-318.

[10] Kamilaris, A. and Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, *147*, pp.70-90.

[11] dos Santos Ferreira, A., Freitas, D.M., Da Silva, G.G., Pistori, H. and Folhes, M.T., 2017. Weed detection in soybean crops using ConvNets. *Computers and Electronics in Agriculture*, *143*, pp.314-324.

[12] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

[13] Mahesh, M. and Kiran, V.K., 2024. Novel image processing architecture using AWS. In *Computational Methods in Science and Technology* (pp. 6-15). CRC Press.

[14] Yang, Y. and Ko, Y.C., 2022. Application of mobile edge computing combined with convolutional neural network deep learning in image analysis. *International Journal of System Assurance Engineering and Management*, *13*(Suppl 3), pp.1186-1195.