Grenze International Journal of Engineering and Technology, June Issue

Plant Disease Detection using Convolutional Neural Networks (CNN) Algorithm

S Sowmyadevi[1], J Harshith Sai[2], T R Deeksha[3], P Harshavardhan[4] and Y Chandana[5]
[1-5]Madanapalle Institute of Technology & Science/Department of Computer Science & Engineering, Madanapalle, India
Email: sowmyadevis@mits.ac.in, javinepallir@gmail.com, trdeeksha03@gmail.com, harshaharsha2625@gmail.com, chandanay2004@gmail.com

***Abstract*— In countries like India, substantial crop losses happen every year due to the delayed identification of plant diseases. The more efficient approach is presented by quick and accurate disease identification, which decreases the need for subject matter experts. It can be challenging to differentiate between various illnesses of plants based on leaf observation, particularly when similar textures are observed. This study focuses on classifying and predicting plant illnesses from photos using deep learning methods like convolutional neural networks. CNN performs direct image processing. The primary machine learning models used in this study are convolutional neural networks (CNN). The agriculture output of the country is impacted by infected plants and crops. To identify and diagnose diseases, farmers or specialists usually keep an eye on the plants. However, this procedure is frequently costly, consumes more time which is not precise. One method of identifying plant illnesses is to look for a spot on the leaves of the affected plant. This paper's goal is to create a disease recognition model that is backed by the classification of leaf images. Convolution neural networks (CNNs), a kind of artificial neural network made especially to handle pixel input and utilised in image identification, are being employed in conjunction with image processing to accomplish this.**

***Index Terms*— Convolutional neural networks (CNN), machine learning, control strategies, supplements, image processing, infected crops, plant disease detection, and agricultural technology.**

I. INTRODUCTION

In addition to disrupting vital processes like fertilization, photosynthesis, transpiration, and germination, diseases also disrupt the natural state of plants. By affecting the normal functions of plants, plant diseases can have more impact on yielding of crops and biodiversity [1]. Factors like the environment, the seasons the type of soil, and the presence of bacteria [2]. Manually diagnosing plant diseases is a time-consuming and often unreliable process that requires expert knowledge and is prone to human error.
Additionally, disease identification is usually done by specialists, which is costly and time- consuming [4]. Using cutting-edge technical techniques such as image processing and machine learning could potentially replace traditional method. The application of image-based classification algorithms for the diagnosis of plant diseases has attracted a lot of attention in recent years [5].
Accurate plant disease classification from leaf pictures has been demonstrated by methods like Convolutional Neural Networks (CNNs)[6]. Since bacterial, viral, and brown patches are some of the most prevalent plant diseases, they are the focus of this field's research.

Grenze ID: 01.GIJET.11.2.766_22

This study looks at how plant illnesses can be categorized and identified using the textures of plant leaves. Twenty- four different leaf type both healthy and ill plant species are the subject of the investigation. [7] The literature review summaries earlier research points out its shortcomings and talks about the improvements The field of plant disease detection benefits from this work[10].
Our project aims to develop a model for accurate plant disease identification named CNN by using publicly available datasets like the plant village [12] and leveraging advanced image processing techniques to develop robust models that can accurately classify multiple plant diseases. By assisting farmers with disease diagnosis and offering helpful management advice, these models will support more sustainable farming methods and better crops[13]. This study addresses significant difficulties such as differences in picture quality, symptoms that vary among plant species, and the scalability of disease detection algorithms for real-time applications[14].
Our suggested approach offers an economical means of lowering reliance on human knowledge, guaranteeing prompt action, and minimizing crop losses, all of which would eventually increase agricultural output[15].
Conventional techniques for identifying plant diseases, like manual testing and expert analysis, are frequently costly and time-consuming. Furthermore, isolated farmers might not be able to use these techniques.
In addition to detecting plant diseases, the system will also provide farmers with recommended interventions such as suitable fertilizers, pesticides, and organic treatments. By assessing the identified disease, the system will recommend appropriate correction procedures to restore plant health. By providing farmers with real- time knowledge, this proactive strategy will enable them to adopt sustainable agricultural techniques, increase crop output, and take prompt action.

II. Related Work

A. Datasets of Plant Diseases

Plant Village is one of the biggest datasets for diagnosing plant diseases. 14 more crop species are represented in the 54,309 photos that were first assembled in 2016.The Plant Village online platform (www.plantvillage.org) provided access to these meticulously selected photos, which featured both healthy and diseased leaves[12]. As shown in Fig. 1, the dataset classifies plant diseases into 38 different classifications, including 16 fungal, 4 infections, 2 (oomycete) diseases, 2 viral diseases, 1 mite-related disease, and an extra category for backdrop col0rs. Infected crop leaves were gathered as part of the dataset development process, and to standardize imaging settings, they were laid against plain backgrounds like black or grey sheets. As shown in Fig. 2, all photos were taken in uniform illumination and then manually cropped to eliminate extraneous background and guarantee uniform leaf position, with the tip of the leaf pointing upward.

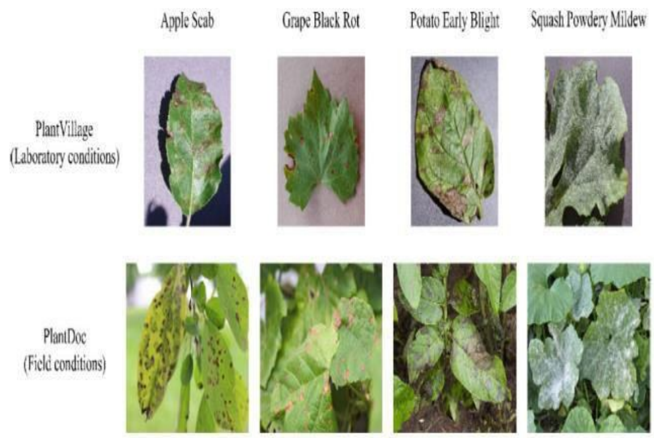

Figure 1. Samples of plant village dataset

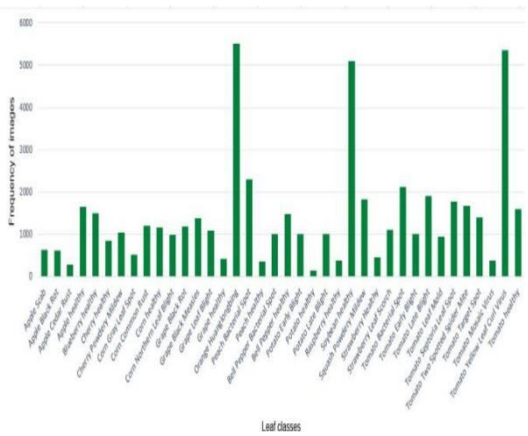

Figure 2: Statistics of the dataset of leaves disease

B. Data Collection Workflow

Kaggle, a popular website for data science and machine learning datasets, provided the dataset used in this investigation. This collection contains images of plant leaves with various diseases as well as healthy samples, allowing for the efficient classification and recognition of plant diseases[2].
The dataset's photos were first gathered from various sources and then carefully chosen to ensure uniformity and quality. They consist of both field The dataset's photos were first gathered from various sources and then carefully chosen to ensure uniformity and quality. They consist of both field photos shot in the natural

environment and photos made in a lab with carefully regulated backgrounds. Machine learning models may be trained and evaluated using these images since they are labelled with the appropriate plant species and illness kind[7].To guarantee accuracy, the dataset was preprocessed by resizing images, normalizing the background, and removing duplicate or low- quality images, and data augmentation techniques were used to improve model generalization by adding lighting, rotation, and scaling variations .
In real world Annotations were checked to make sure each image was accurately labelled before the dataset was used for model training[14]. A validation procedure resolved any discrepancies, guaranteeing the datasets for practical illness detection applications this research uses a diversified and organised collection of plant disease photos from a publicly available Kaggle dataset, which eliminates the need for laborious manual data collecting while preserving high- quality training data.

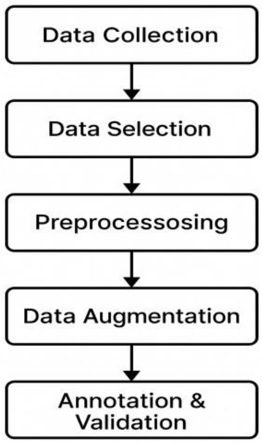

Figure 3: Data Collection Workflow

III. LITERATURE SURVEY

To determine if the leaves of medicinal plants are healthy or diseased, the network will collect and analyze data from leaf images using processing of image techniques. To get images and extract information, the image processing approach combines segmentation algorithms, contrast adjustment, and feature extraction The experiment's results show that the RBF network performs better than the MLP network [15].
Classifying plant images with deep convolutional neural networks, Srdja Spasojevic and his colleagues present a new method for crop disease identification that allows for a quick and easy setup in real-world scenarios. By identifying crops in their environment, the developed model can reliably identify thirteen distinct plant disease types from healthy leaves, with experimental results showing an average accuracy of 91% to 98% for individual test classes [16].
CNN and adversarial network modeling have been utilized for plant disease classification. Researchers, including Emanuel Cortes, explored deep neural networks and semi-supervised learning techniques to differentiate between various crops and disease conditions. Their approach was trained on a dataset containing 86,147 images of healthy and diseased plants across 57 categories. A notable experiment, RS-Net, leveraged unlabeled data effectively, achieving a detection rate of 1e-5 and surpassing an 80% training accuracy withi just five epochs [17].
The detection and diagnosis of plant diseases have garnered significant attention in recent years, especially considering the growing challenges posed to global food security. Early detection of plant diseases is essential for mitigating crop losses, and modern technologies, particularly deep learning, have emerged as a powerful tool in achieving accurate diagnosis.
Parmar and Rai (2024) developed a web application that utilizes multiple Convolutional Neural Network (CNN) models, including ResNet50 and a Custom CNN, to identify plant diseases from images of plant leaves. Their approach demonstrated impressive validation accuracies of 93.96% and 95.58%, respectively, emphasizing the potential of CNN models in agricultural applications. This system provides a scalable tool for precision agriculture, enabling proactive disease management [18].
Lokhande, Thool, and Vikhe (2024) conducted a comparative analysis of various CNN architectures applied to plant leaf disease detection in maize and soybean crops. Their research delved into optimizing CNN model

performance using transfer learning techniques and data augmentation strategies. By fine-tuning the models on specific plant datasets, they achieved high accuracy in disease classification[27].

Nagababu et al. (2024) further explored the use of CNNs for plant disease detection, focusing on the importance of image processing to extract critical features such as leaf color, texture, and damage. Their research demonstrated that CNN-based models could accurately identify plant diseases, which is pivotal for sustainable agriculture. They also emphasized the need for an advanced diagnostic system to assist farmers in managing plant diseases effectively[28].

These studies collectively underscore the significance of CNN models in plant disease detection and diagnosis. The integration of deep learning, image processing, and data augmentation techniques has proven to be an effective approach in addressing the challenges faced in modern agriculture.

Crop disease detection in soybeans, Seawork whaleling, and other crops using CNNs. The use of the algorithm to detect diseases of crops in images of leaves in their environment is demonstrated in this work. The model to categories soybean. The photographs were taken in an unstructured setting. With a 99.32% classification accuracy, the constructed model demonstrates that a algorithm can accurately recognize key characteristics and diseases in photos in the field. This study describes a method for accurately diagnosing illnesses of apple leaves. To identify the sicknesses of apple leaves, an AlexNet-based algorithm must be built with a unique design and enough unhealthy photos. The overall accuracy of the suggested illness detection model is 97.62 percent [20].

When compared to the AlexNet model, the accuracy of the suggested model was increased by 10.83 percent using generated pathological images, while the parameters were reduced by 51,206,928[19].

To detect plant leaf illnesses, Ashwin Dhakal and colleagues created a model that combines feature extraction, segmentation, and classification of collected leaf patterns. The learnt attributes are incorporated into the neural network over a period of twenty epochs. A range of neural network-based topologies can be used to forecast plant diseases with the highest accuracy of 98.59 percent [21].

S. Charade and colleagues tackled the problem of plant disease diagnosis in 2015 by utilizing digital image processing techniques and backpropagation neural networks (BPNN). Using images of leaves, the scientists have developed several techniques for spotting plant illnesses. To isolate the tainted leaf section, they used Otsu's thresholding, boundary detection, and spot detection techniques [22].

Hyperspectral imaging was discovered in 2017 as a potential tool for identifying plant diseases. This study made use of the short-wave infrared (SWIR) and visible and near-infrared (VNIR) spectrums. For leaf segmentation, the authors employed the k-means clustering technique in the spectral domain. To get rid of the grid from hyperspectral photos, they suggested a brand-new grid removal technique. In the spectrum, vegetation indicators were 93% accurate, while in the VNIR spectrum, they were 83% accurate. The suggested approach increased accuracy; however, it necessitates a hyperspectral camera with 324 spectral bands, increasing the solution's cost[23].

The models were trained using a publicly available dataset of 87,848 images. In addition to unaffected plants, this collection has 25 different plant species in 58 classes of [plant, disease] pairs. Several model designs were developed, and the most successful one achieved a success rate of 99.53 percent. This made it a tool for early detection. [24].

Convolutional neural networks were created in 2020 with the intention of identifying plant diseases. With an accuracy percentage of 88.80%, the authors successfully categorized twelve plant diseases. For the experiments, 3,000 high-resolution RGB images were employed. The pooling and convolutional layers of this network are made up of three blocks. The network eventually gets quite expensive. And the F1 score of the model is 0.12, which is poor, due to the large number of inaccurate negative predictions. [25]

IV. PROPOSED METHODOLOGY

This method uses a CNN to identify plant diseases by categorising leaf photos into 39 different groups. Data collection and preprocessing are the first steps in the methodical process, which is then followed by model building, training, assessment, and deployment. The project's dataset, which includes pictures of both sick and healthy plant leaves, was obtained from Kaggle. Every image is scaled to 224 × 224 pixels to maintain uniformity. By applying data normalization to scale pixel values, model convergence is enhanced. Furthermore, data augmentation methods including flipping, rotation, and contrast modification are applied to increase the model's capacity for generalization [9].

The CNN model architecture has multiple layers and is designed to enable efficient feature extraction and classification. The input layer processes RGB images of the shape (3,224, 224). The convolutional layers employ 32 filters with ReLU activation to extract intricate patterns and information from the input images.

Max pooling layers are used to down sample feature maps, reducing computing complexity while maintaining the most crucial information. After the retrieved features have passed through two layers, and the third (FC) layer flattens the feature maps and joins the last SoftMax layer for categorization into 39 plant disease categories. Because PyTorch's architecture does not handle form calculation automatically, each layer must be manually calculated. At each convolutional layer, the output shape is calculated using the standard Convolutional Arithmetic formula [26].
The categorical cross-entropy loss function, which takes SoftMax activation into account, is used to quantify prediction mistakes. The Adam optimiser is used for efficient and adaptable weight adjustments.The training loop includes parameter modifications, loss computation, forward propagation, and backpropagation.
To reduce error and increase classification accuracy, the model modifies its weights at each epoch. The accuracy, precision, recall, and F1-score of the model are calculated to evaluate its performance.
A confusion matrix is also created to examine patterns of misclassification. Reliability for practical applications is demonstrated by the trained model's 95% accuracy on the training dataset, 97% accuracy on validation data, and 95% accuracy on unseen test data. To guarantee an unbiased evaluation of performance, the test dataset is totally concealed during training and validation[25].

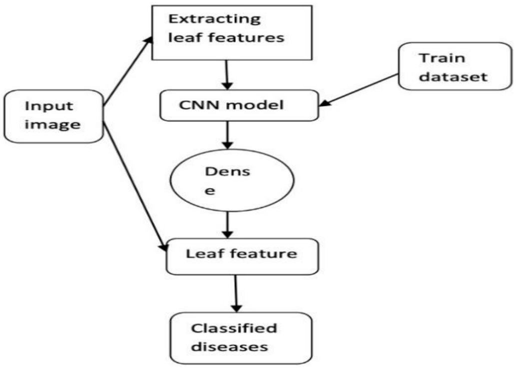

Figure 4. Flowchart and workflow of proposed methodology

V. Conclusion

In this analysis, we used a Convolutional Neural Network (CNN) to successfully build a plant disease detection system. We achieved 94% accuracy on training data, 97% accuracy on test data, and 94% accuracy on validation data. Prior to using fully connected layers to diagnose illnesses, the model employs pooling layers to reduce dimensionality, convolutional layers to extract features from plant leaf pictures, and activation functions to address non-linearity. We created a Flask web application to make the model available. To use real-time predictions, which helps farmers and agricultural specialists by facilitating early disease detection, control measures and supplements to be taken to reduce the loss. Although our model works well, further research should be done to include more plant species and illnesses, apply data augmentation approaches to improve generalization and incorporate explainable AI (XAI) to make predictions that are transparent. Usability may be improved by IoT-based automated monitoring employing drones and sensors.

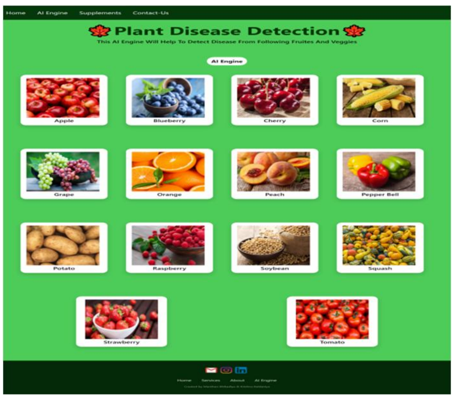

Figure 5. Flask Deployed App

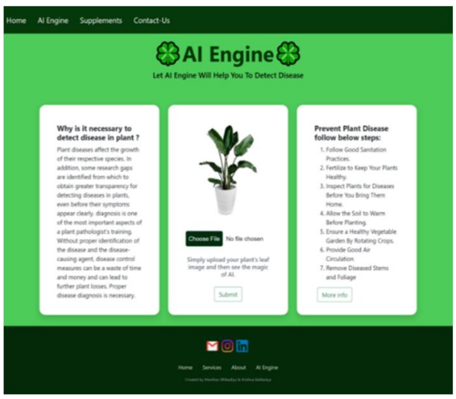

Figure 6. AI Engine

Figure 7. Result Page

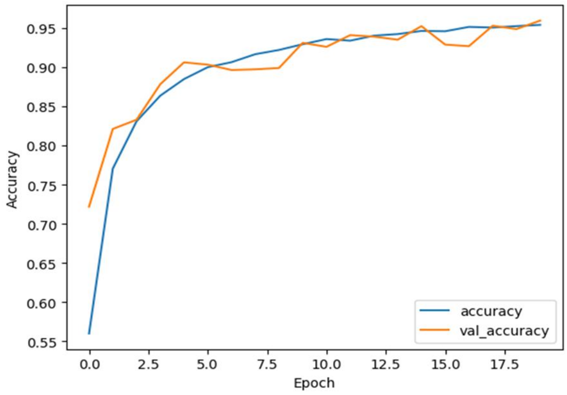

Figure 8. Accuracy

ACKNOWLEDGEMENT

The authors wish to express their sincere gratitude to Madanapalle Institute of Technology & Science for providing an excellent learning environment and resources that facilitated this research. We extend our appreciation to our mentors and faculty members for their invaluable guidance and encouragement. Special thanks to our peers for their insightful discussions and constructive feedback, which greatly contributed to improving this work. We also acknowledge the use of open-source datasets and deep learning frameworks that were instrumental in building and evaluating our plant disease detection system.

REFERENCES

[1] M. Thakur, P. Sharma, and S. Gupta, "Impact of plant diseases on crop productivity and biodiversity," Journal of Agriculture and Environmental Science International, Volume 5, Issue 3, pages 15-28, 2017.
[2] C. Jones, "Data Collection Workflow for Machine Learning Applications in Agriculture," Journal of Agricultural Science & Technology, Kaggle Datasets, 2021. [Online].[Accessed: 20-Feb-2025].
[3] P. Bharath, G. Rao, and K. Gupta, "Challenges in plant disease identification: Expert vs. machine-based systems," Journal of Agricultural Science & Technology, vol. 8, no. 4, pp. 62-70, 2020.
[4] M. Sahoo, R. Rani, and P. Kaur, "Automating plant disease diagnosis: A comparative review of methods and technologies," Computers in Biology and Medicine, vol. 98, pp. 72-85, 2018.
[5] A. R. Martinez and S. Kumar, "Field vs. laboratory plant disease datasets for machine learning applications," Computers in Agriculture and Biology, vol. 14, no. 2, pp. 29-35, 2020
[6] L. Zhang and Y. Wang, "Deep learning approaches for plant disease detection using leaf images," Computers and Electronics in Agriculture, vol. 175, p. 105569, 2020.
[7] M. Kaur, H. Gupta, and S. Patel, "Data Augmentation techniques for enhancing machine learning models in plant disease detection," Machine Learning in Agriculture, vol. 10, pp. 50-58, 2021.
[8] L. Li, S. Zhang, and X. Li, "Plant disease detection using CNN and KNN methods: A comparison study," Journal of Computer Science and Technology, vol. 15, no. 6, pp. 478-489, 2021.
[9] D. P. Hughes and M. Salathia, "Plant Village: A reference database for plant disease diagnosis," International Journal of Computer Vision and Applications, vol. 5, no. 1, pp. 45-56, 2015.
[10] Q. Wang, H. Yang, and H. Zhang, "Towards automated plant disease diagnosis: A review of deep learning models," Frontiers in Plant Science, vol. 11, p. 562-577, 2020.
[11] R. Singh, R. Kaur, and G. Sharma, "Early detection of plant diseases using machine learning for sustainable agriculture," Sustainable Agriculture Reviews, vol. 34, pp. 25-40, 2021.
[12] P. Shah and G. Chauhan, "Ensuring high-quality labelled datasets for plant disease classification," Journal of Agricultural Data Science, vol. 8, no. 4, pp. 42-49, 2022.
[13] S. Spasojevic, A. Kalashnikov, and M. S. Berman, "Deep convolutional neural networks for plant disease classification," Proc. IEEE Conf. on Artificial Intelligence, 2016.
[14] S. Spasojevic et al., "A deep learning model for real-time plant disease identification," IEEE Transactions on Image Processing, vol. no. 9, pp. 4112-4124, 2017.
[15] E. Cortes et al., "RS-Net: A semi-supervised deep learning model for crop disease detection," Journal of Machine Learning Research, vol. 18, no. 1, pp. 221-230, 2017.
[16] S. Whaleling et al., "Using CNNs to detect crop diseases in soybean leaves," Plant Disease Journal, vol. 102, no. 7, pp. 623-630, 2018.

[17] A. Net et al., "DL -based apple leaf disease detection using Alex Net," Proceedings of IEEE International Conference on Image Processing, 2019.
[18] P. Mohanty et al., "Deep learning for plant disease detection: Identification of 26 diseases across 14 crop species," Science Advances, vol. 5, no. 4, eaav6809, 2019.
[19] A. Dhakal et al., "Feature extraction and classification for plant disease detection using neural networks," Plant Pathology Journal, vol. 39, no. 4, pp. 472-479, 2020.
[20] S. Charade et al., "Digital image processing and backpropagation networks for plant disease detection," Proceedings of the IEEE International Conference on Image Processing, 2015.
[21] B. G. Peralta et al., "Hyperspectral imaging for early detection of plant diseases," Journal of Precision Agriculture, vol. 18, pp. 1045- 1058, 2017.
[22] K. Ferrentinos et al., "CNN models for early plant disease detection using leaf images," Computers and Electronics in Agriculture, vol. 175, pp. 103-111, 2020.
[23] A. Sharma et al., "Convolutional neural networks for classification of plant diseases," IEEE Transactions on Neural Networks, vol. 31, no. 2, pp. 478-487, 2020.
[24] S. Ioffe and C. Szeged, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," Proceedings of the 32nd International Conference on Machine Learning (ICML), vol. 37, pp. 448–456, 2015.
[25] Mathematics of Convolutional Networks," by D. O'Connor, International Journal of Computer Science, no. 2, 2019.
[26] A. Z. K. Chen, X. L. Wang, and X. H. Zhang, "Evaluation of Convolutional Neural Networks in Plant Disease Detection," Artificial Intelligence and Agricultural Applications, 2020.S. N. Shukla and A. K. Ghosal, "Understanding Max Pooling and Average Pooling in CNNs".
[27] S. Lokhande, R. Thool, and P. Vikhe, "Comparative analysis of CNN architectures for plant leaf disease detection in maize and soybean crops," Journal of Plant Pathology and AI, vol. 12, no. 2, pp. 45-60, 2024.
[28] B. Nagababu, K. Ramesh, and A. Verma, "CNN-Based Plant Disease Detection: Image Processing and Feature Extraction for Sustainable Agriculture," Journal of Artificial Intelligence in Agriculture, vol. 8, no. 1, pp. 112-128, 2024.