2024 IEEE 9th International Conference for Convergence in Technology (I2CT)
Pune, India. Apr 5-7, 2024

Beyond 2D: 3D Point Cloud Analysis with Color Extraction and Precise Percentage Calculation of Anthracnose in Chillies

Sneha Varur
School of CSE
KLE Technological University
Hubballi, India
sneha.varur@kletech.ac.in

G R Sampreeti
School of CSE
KLE Technological University
Hubballi, India
sampreetigr25@gmail.com

Sushma Bhat
School of CSE
KLE Technological University
Hubballi, India
bolmanesushma@gmail.com

Vaishnavi J Ajjevadeyarmath
School of CSE
KLE Technological University
Hubballi, India
ajjevadeyarmathvaishnavi@gmail.com

Shweta Pagadi
School of CSE
KLE Technological University
Hubballi, India
shwetapagadi2021@gmail.com

Meena S M
School of CSE
KLE Technological University
Hubballi, India
msm@kletech.ac.in

2024 IEEE 9th International Conference for Convergence in Technology (I2CT) | 979-8-3503-9447-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/I2CT61223.2024.10544162

***Abstract*—This paper introduces a novel Deep Learning approach leveraging the ResNet50 Architecture for the classification of diseased red chili peppers. Beyond conventional 2D methods, the proposed methodology incorporates a 3D point cloud approach to accurately determine the percentage of rotten chili, aiming to revolutionize crop health management. The comprehensive process includes data collection, 2D image classification, 2D to 3D point cloud conversion, and defected region extraction, applied to a meticulously assembled dataset. ResNet-50 proves effective, and the integration of a color exclusion algorithm enhances the analysis of diseased regions, providing a quantitative metric for evaluating filtering processes. This methodology significantly contributes to the resilience of chili crop production and holds promise for broader applications in plant disease detection, emphasizing the importance of integrating 3D point cloud methodologies in agricultural research and technology.**

***Index Terms*—Red chilli, Anthracnose, Convolutional Neural Networks(CNN), ResNet50, 3D Point Cloud, Diseased Region Extraction,Severity Calculation**

I. Introduction

Red chili, scientifically known as Capsicum annuum, holds immense global significance as a spice embedded in diverse culinary traditions. In this landscape, India, often referred to as the ”Land of Spices,” plays a crucial role in the worldwide spice trade, leaving a profound impact on culture, economics, and agriculture [1]. The influence of India's connection to red chillies extends beyond culinary heritage, shaping global tastes and economies. Additionally, red chillies boast more Vitamin A than carrots and contain 'Capsaicin,' an active component providing antioxidants and other health benefits.However, despite its importance, red chili faces significant threats, notably 'Chilli Anthracnose,' caused by the Colletotrichum fungus, posing a global risk to crops [2].

Early detection of chili diseases is vital, leading to the widespread use of image processing techniques in agriculture. Deep Learning and Machine Learning techniques, including the Random Forest Classifier, Multiple Linear Regression, Genetic Algorithms, and Convolutional Neural Networks (CNN) [3], contribute to identifying plant diseases. The Random Forest Classifier, for instance, leverages various image features to make decisions about disease presence, while Multiple Linear Regression establishes correlations between image features and the likelihood of a plant being diseased.CNN architectures like MobileNet, DenseNet, ResNet, and others are employed for precise plant disease identification [4]. Image segmentation using K-means clustering and the Support Vector Machine (SVM) method further aid in plant disease detection [5].

Our paper introduces a novel Deep Learning solution utilizing the ResNet50 Architecture for the classification of diseased chili peppers. While traditional 2D classification methods exist, they fall short in quantifying the extent of chili damage. Our study introduces a 3D point cloud methodology designed to determine the percentage of rotted chili. The integration of this methodology with Convolutional Neural Networks (CNN) enhances the assessment of red chili plants, promising valuable insights for farmers in making informed decisions for sustainable and efficient crop production.

This technological advancement not only improves disease detection but also provides a more profound understanding of the spatial dynamics of chili plants. By bridging the gap between 2D classification and 3D point cloud methodologies, our proposed solution has the potential to revolutionize crop health management. In the face of global challenges like Chilli Anthracnose, this innovative approach emerges as a beacon for farmers, contributing to the resilience and sustainability of chili crop production and holding promise for broader applications in plant disease detection across various crops.

II. Background Study

In the field of image processing, Convolutional Neural Networks (CNNs), including ResNet-50, are pivotal for proficient feature extraction and image classification. There's a rising interest in converting 2D images into 3D point clouds, and the architecture referred to as Base40M plays a critical role in this transformation. This investigation aims to evaluate the efficacy of CNNs, particularly ResNet-50, in adeptly extracting features and classifying images, ultimately contributing to the conversion of images into 3D point clouds.

A. Image Classification:

1) CNN: Convolutional Neural Networks (CNNs) excel in 2D image classification [6], [7], [8] by employing a series of mathematical operations within their layers. The convolutional layers in Fig 1, characterized by the convolution operation

$$(I * K)(i,j) = \sum_{m,n} I(m,n)K(i-m, j-n) \quad (1)$$

and apply filters (kernels) to extract features from the input image. Introducing non-linearity through an activation function, typically ReLU, enhances the network's ability to capture intricate patterns. Subsequently, max pooling layers, employing the operation reduce spatial dimensions by selecting the maximum value from local regions, thereby preserving essential features. The fully connected layers consolidate these features through a weighted sum

$$z = \sum_i w_i \cdot x_i + b \quad (2)$$

and an activation function, leading to higher-level representations. Finally, the softmax activation function

$$P(class_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (3)$$

in the output layer transforms raw scores into class probabilities, enabling the network to make informed predictions. The entire classification process is encapsulated by the formula

$$\text{Class} = argmax\_(i) * P(\text{class}\_i) \quad (4)$$

where the class with the highest probability is identified as the ultimate classification. This intricate interplay of formulas highlights how CNNs systematically extract features and classify images, making them powerful tools in the field of computer vision.Fig 1. shows the general architecture of a CNN classification.

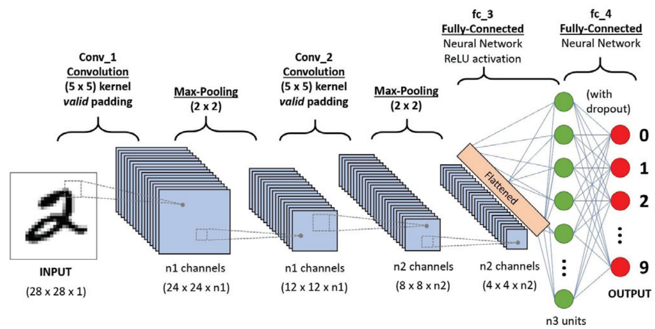

Fig. 1. CNN 2D Classification Architecture

2) ResNet Model: In ResNet (Residual Network) for 2D image classification [9], [10], [11], [12] the key idea is to simplify the learning process for deep networks. Instead of directly trying to learn the mapping for each layer, ResNet uses ”residual blocks” with a formula like:

$$Output = Input + LearnedTransformation \quad (5)$$

the network learns to adjust the input rather than transforming it completely. This helps in avoiding potential issues with very deep networks. The overall architecture stacks these blocks together. For instance, a basic block looks like:

$$Output = Activation(Convolution(Input)) + Input \quad (6)$$

The addition of the input to the transformed output helps in the smooth flow of information during training, making it easier to train deep models effectively.

B. 2D Image To 3D Point Cloud Conversion:

A collection of data points in three-dimensional space is referred to as a 3D point cloud. Each point within this set is characterized by its X, Y, and Z coordinates.It serves as a comprehensive representation of surfaces and structures in a 3D environment, capturing spatial details.Generated through diverse 3D sensing technologies, each point corresponds to a precise location in physical space [13].

The use of point clouds has become crucial in representing three-dimensional spaces, with applications spanning robotics, autonomous driving, and augmented and virtual reality [14]. Deep learning, a dominant force in the field of computer vision, is widely favored for tasks such as classification, segmentation, and detection.

The rising prevalence of 3D acquisition technologies, including LiDARs, 3D scanners, and RGB-D cameras, has heightened the need for effective methods to convert 2D images into 3D point clouds [15].An alternative technique for rapid 3D object generation involves the utilization of the Base-40M model. This approach enables the creation of 3D models within a brief period of 1-2 minutes using a single GPU. The process generates a diffusion model to produce a 3D point cloud [20].

III. Related Work

This section reviews recent advancements in automating chilli disease classification using 3D point clouds. Recognizing the significance of chilli as a vital Indian crop prone to diseases causing yield losses, we emphasize the need for early and accurate detection. Traditional diagnostic methods are acknowledged for their time and cost challenges, prompting exploration into automated methods leveraging 3D point clouds to capture intricate features of chilli plants [16].

A. Deep Learning Architectures for Image Recognition:

The author, Liang et al., conducted a comprehensive exploration of challenges in deep learning models for image recognition [17]. They introduced innovative architectures such as ResNet, Inception-ResNet, and Inception-ResNet v1, which incorporated residual connections and inception modules, showcasing unparalleled accuracy in image identification. The investigation aims to apply these architectures to advance 3D point cloud classification, specifically in the context of chilli disease.

B. CNN Mastery: Transforming Image Classification with Impressive Results:

Sharma and Phonsa demonstrated the dominance of Convolutional Neural Networks (CNNs) in image classification, achieving 94% accuracy on the CIFAR-10 dataset for classes like aeroplanes, birds, and cars [18]. CNNs, implemented in a Jupyter Notebook, showed remarkable efficiency with a 90% validation accuracy after just 20 epochs. The study suggests that the efficacy of CNN techniques varies with project complexity, with multi-label approaches often outperforming alternative models, highlighting CNNs' potential to revolutionize image classification.

C. PointE: Text-Driven 3D Cloud Innovation:

The PointE system by Nichol et al. pioneers 3D point cloud synthesis using a Transformer-based architecture, Gaussian noise, CLIP-based clustering, and heuristics [19]. Despite resolution constraints, PointE excels in generating intricate shapes, holding promise for high-fidelity 3D representations with implications for real-world fabrication. Inspired by advancements like ResNet, PointE significantly contributes to the evolving landscape of 3D synthesis methodologies.

D. 3D Point Cloud Generation from 2D Images:

Chen et al. introduced a cutting-edge point cloud generation network that integrates image cropping, retrieval, and point cloud reconstruction [20]. Leveraging VGG-16's feature extraction capabilities, this network adeptly weaves intricate point cloud models with precision and accuracy. Their exploration into multi-task architectures for 3D point cloud classification is poised to enhance understanding of complex 3D environments, opening new avenues for advanced applications in various domains.

E. Structure from Motion (SFM) for 3D Point Cloud Generation:

Zhang et al. applied Structure from Motion (SFM) for generating immersive 3D views from stereo and monocular images [21]. Their work contributes to creating realistic and highly detailed 3D views, which can be converted into point clouds. This research holds pivotal significance in point cloud classification, particularly in the identification and categorization of chilli disease. By enriching point cloud data with realistic representations, Zhang and team's work serves as a cornerstone for advancing methodologies in chilli disease detection through 3D reconstructions.

IV. Proposed Methodology

A. Data Collection:

Fig. 2. Samples of Dataset of Red chilli both front and back view Healthy, Partially Diseased and Fully Diseased

Due to the limited availability and absence of pertinent data, we undertook the task of assembling and constructing our dataset. This dataset was created to meet our specific needs for integration into the Deep Neural Network (DNN) architecture, designed for the classification, 3D Point cloud generation and diseased region extraction of Anthracnose-affected chillies and healthy ones.

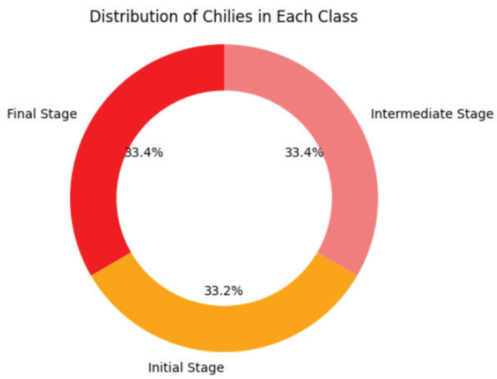

Fig. 3. Graphical representation of chilli dataset

To compile this dataset, we gathered chilli samples from two primary sources: the University of Agricultural Sciences in Dharwad and chilli farms located near Kusugal on the outskirts of Hubli. Our selection included both healthy and disease-affected chillies. The University's collaboration ensured that we obtained chillies exclusively afflicted with Anthracnose, eliminating confusion with other diseases.

The collected samples comprised Red Chillies only. The images were captured under controlled conditions against a white backdrop in a well-lit room. Utilizing a high-resolution camera mounted on a tripod, we took equidistant pictures, maintaining an aspect ratio of 1:1. Each chilli's both front view and back view images were captured.This meticulous process aimed to ensure the quality and accuracy of our dataset for subsequent use in our project.Samples of the images captured are shown in Fig 2.The 'donut graph' in Fig 3 shows the percentage of images in each type.

B. 2D Image Classification:

ResNet-50's prowess in chili image classification stems from its deep convolutional neural network (CNN) architecture. Trained on diverse datasets, the model excels at feature extraction, utilizing its layers as specialized filters.

For moderately affected chilies, ResNet-50's deep layers meticulously scrutinize images, capturing subtle disease indicators for precise early detection.In cases of severely affected chilies, the model's depth becomes crucial, discerning complex patterns and severe symptoms

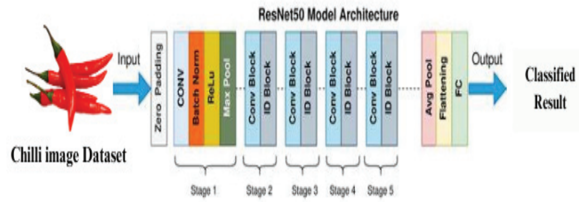

Fig. 4. ResNet-50 Architecture

indicative of advanced disease stages. The hierarchical feature representation allows ResNet-50 to accurately identify nuances associated with different disease severities.

In the case of healthy chilies, ResNet-50's negative feature recognition plays a key role. Trained to recognize the absence of disease-related patterns, the model ensures accurate classification by identifying distinctive features associated with healthy plants. This technical proficiency, grounded in a deep CNN framework, empowers ResNet-50 to distinguish between varying degrees of disease severity and reliably classify chili images. The model's intricate understanding of visual cues contributes to its effectiveness in aiding farmers and researchers in precise disease management and crop health improvement.Fig 4. depicts a ResNet-50 architecture.

C. 2D Image To 3D Point Cloud Conversion:

The background study encourages the utilization of the base40M model, a diffusion model comprising 40 million parameters specialized for converting images to point clouds. This model is dependent on the latent grid obtained from a CLIP ViT-L/14 model, where CLIP stands for Contrastive Language-Image Pre-training. The notation ViT-L/14 is likely indicative of a particular version of the Vision Transformer (ViT) model with 14 layers.

The point clouds generated by this model represent three-dimensional data points, each characterized by (x, y, z) coordinates and (R, G, B) colors, all normalized within the range [-1, 1]. The generation process employs a diffusion technique, starting from random noise and gradually refining it over time.The noising process is defined as

$$q(x_t|x_t-1) := N(x_t; \sqrt{1-\beta_t x_t - 1}, \beta_t I) \tag{7}$$

Where β_t is a schedule determining the amount of noise added at each time step.

The model architecture is Transformer-based (inspired by Vaswani et al., 2017) and predicts both ε and $\sum$ parameters conditioned on the input image, timestep (t), and the noised point cloud (x_t). Each point in the point

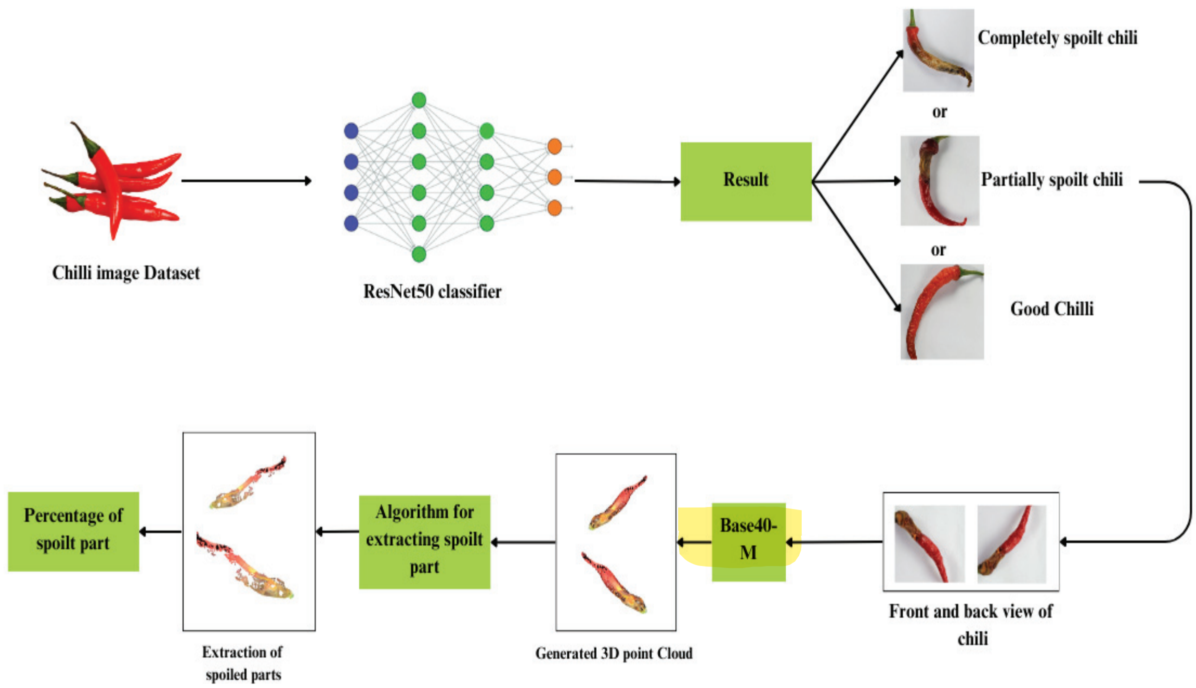

Fig. 5. Architecture Design Of Our Model

cloud undergoes linear transformation through a layer. The timestep is processed through a small Multilayer Perceptron (MLP). Additionally, the image is processed using a pre-trained ViT-L/14 CLIP model, and its last layer embeddings are linearly projected to align with the model context. The final output is a sequence of predictions for ε and $\sum$corresponding to the input points in the point cloud.

The base model generates a low-resolution point cloud for resolution enhancement Point Cloud Upsampler is employed. The Upsampler shares the architecture of the base model but includes extra conditioning tokens specific to the low-resolution point cloud. This mechanism is utilized to augment the resolution of the generated point cloud.

Conventional 2D images capture only the surface features of plants, making it challenging to detect subtle symptoms of diseases like anthracnose. To overcome this limitation, we employ a method that transforms 2D images into 3D point clouds, enhancing the depth and information layers. Utilizing this approach, we successfully identify and quantify the proportion of spoiled parts. The algorithm processes point cloud data, considering both the front and back perspectives of a chili. It selectively eliminates specific colors within a predefined tolerance, representing the damaged areas. Subsequently, the code calculates the percentage of retained points, offering a quantitative measure to evaluate the effectiveness of the filtration process. This analysis is designed to explore how the exclusion of certain colors impacts the overall structure of the chili's point cloud, providing valuable insights into the preservation of structural details in both views.Fig 5. depicts the architecture and flow of how our proposed model works.

D. Diseased Region Extraction And Percentage Calculation:

In response to the challenges inherent in accurately assessing chili rot, particularly when relying on 2D image analysis, our algorithm introduces a pioneering solution.This methodology involves the processing of point cloud data, capturing both the frontal and dorsal perspectives of a chili. Through a targeted exclusion of specific colors within a defined tolerance, the algorithm identifies a distinct subset of points that persist post color-based filtering. Utilizing code, the algorithm quantifies the percentage of retained points, providing a quantitative measure of the filtering process's effectiveness. The overarching analysis seeks to understand the impact of color exclusion on the overall composition of the chili's point cloud, shedding light on the preservation of structural details in both views.

ALGORITHM: Calculating Percentage Of Diseased Region

Input:

- Generated 3D point cloud of the front view of a chilli $\rightarrow \boldsymbol{F}$
- Generated 3D point cloud of the back view of a chilli $\rightarrow \boldsymbol{B}$

// Each point is of dimension (x,y,z) coordinates and (R,G,B) colors.

Output:

- Percentage of the diseased region $\leftarrow \boldsymbol{P}$

1. ***Initialize*** color_tolerance.
2. Calculate the number of points in $\boldsymbol{F} \rightarrow \boldsymbol{T_{len}}$.
3. Select the 3D point cloud of the front view of the chilli $\boldsymbol{F}$.
4. ***for*** each point in $\boldsymbol{F}$ **do**
   if color is neither Black nor Red nor Green ***then***
   S=color segmentation(F)
   end if
   end for

// Add points if absolute difference between the color of a point and the colors(Red, Green,Black) exceeds the specified color_tolerance.

5. Count the number of points in $\boldsymbol{S}$.
6. ***Repeat*** steps 3, 4, and 5 for $\boldsymbol{B}$.
7. Add the number of points in both the front and back sides $\rightarrow \boldsymbol{D_{len}}$.
8. Calculate the percentage:

$$\boldsymbol{P} = \frac{\boldsymbol{D_{len}}}{\boldsymbol{T_{len}}} \times 100$$

9. Print $\boldsymbol{P}$.

Fig. 6. Algorithm for defective region 3D point cloud extraction and percentage calculation

V. RESULTS

This study presents a methodology for the extraction and computation of the diseased region percentage in red chilli. Initially, we employ the ResNet-50 model to classify the 2D chilli images. After 15 epochs, the model achieves Validation accuracy of 90.67%. The accuracy and validation accuracy for the initial 5 epochs are detailed in Table 1, with corresponding accuracy and loss plotted in Fig. 7 and Fig. 8.

TABLE I
RESNET-50'S ACCURACY AND VALIDATION ACCURACY TABLE FOR 5 EPOCHS

Epochs	Accuracy	Validation Accuracy
Epoch-1/15	0.9353	0.7833
Epoch-2/15	0.9374	0.8583
Epoch-3/15	0.9708	0.8750
Epoch-4/15	0.9916	0.8917
Epoch-5/15	0.9896	0.8750

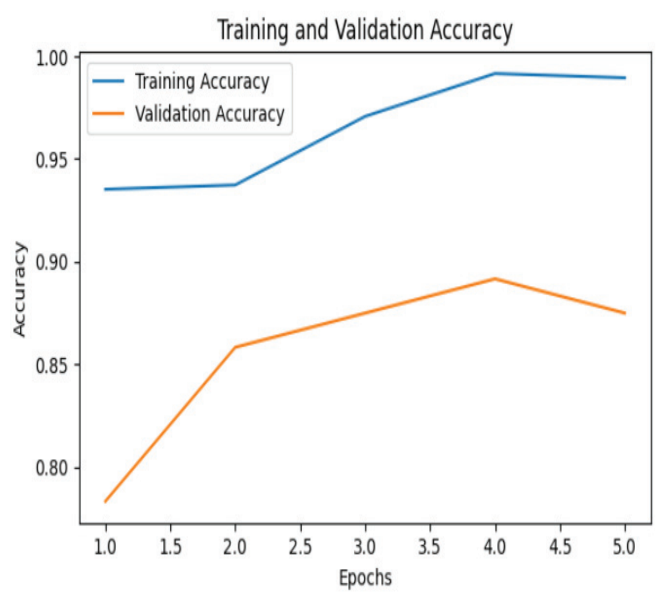

Fig. 7. ResNet-50 Accuracy Graph

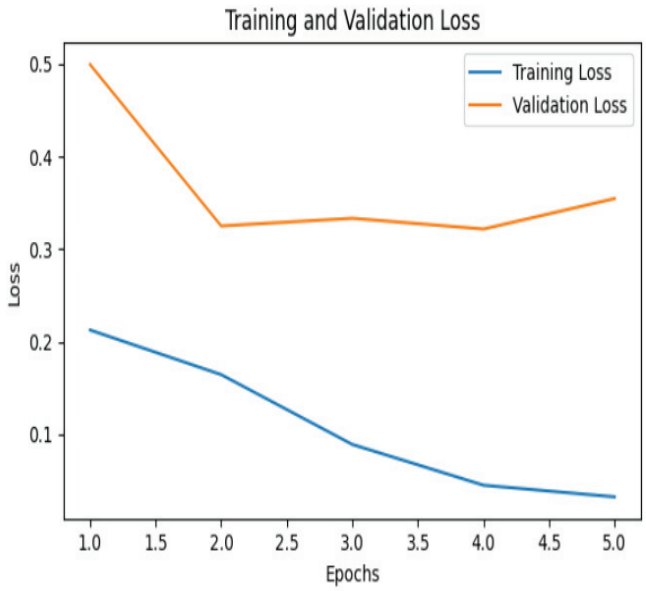

Fig. 8. ResNet-50 Loss Graph

Subsequently, the 2D chilli image undergoes transformation into a 3D Point Cloud using the Base-40M model.This 3D point cloud becomes instrumental in the subsequent extraction process.In the final phase, our custom algorithm is employed for the extraction of the 3D point cloud representing only the diseased region of the chilli. Simultaneously, the algorithm calculates the percentage of the defective portion, providing valuable insights into the extent of the diseased area.Fig 9. depicts the image of a red chilli, it's 3d point cloud and extracted point cloud.

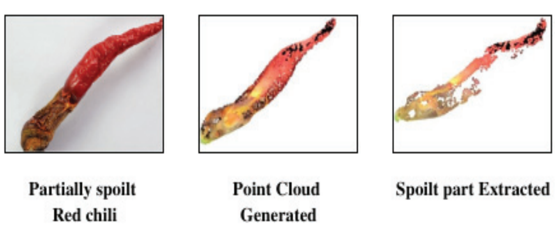

Fig. 9. 2D image of a red chilli, it's 3D point cloud and extracted 3D point cloud images

VI. Conclusion And Future Scope

Chili, as a crucial spice globally, encounters the challenge of various diseases that can significantly impact its quality. Our proposed method, leveraging the ResNet50 architecture, provides an effective means of classifying chili as either good, partially spoilt, or completely spoilt. Additionally, we converted 2D images to 3D point clouds and determined the extent of chili damage through the "Calculating Percentage of Diseased Region" algorithm. Our method has demonstrated efficacy, exhibiting a reasonable level of accuracy, and it serves as a foundation that can be further refined and optimized. There is considerable potential for the advancement and fine-tuning of our current model.The color tolerance specification currently plays a pivotal role in accurately calculating the affected percentage of chili. However, moving forward, there is an opportunity to explore and discover a more optimal solution that eliminates the necessity for setting specific tolerance values. Beyond immediate enhancements, future research could explore real-time monitoring systems and advanced data analytics. This exploration could contribute to the practical implementation of our proposed solution in agricultural settings, providing farmers with timely and precise information for proactive crop management decisions.

References

[1] Bagal, N.U., Kshirsagar, P.J., Torane, S.R. and Manerikar, S.S. (2020). Export of spices from India: An instability analysis. Internat. Res. J. Agric. Eco. & Stat., 11 (1) : 55-63, DOI : 10.15740/

[2] Raj Kiran, Jameel Akhtar, Pardeep Kumar, and Meena Shekhar. Anthracnose of chilli: Status, diagnosis, and management.In Capsicum.

[3] Abou El-Maged, M Lobna, Ashraf Darwish, and Aboul Ella Hassanien. Artificial intelligence-based plant diseases classification. In Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges, pages 45–61. Springer, 2021.

[4] Santoshreddy Machenahalli, VB Nargund, and Suresh Patil. Quick detection and diagnosis of chilli fruit rot pathogens. Vegetos, 27(3):188–191, 2014.

[5] Abdul Hafiz Bin Abdul Wahab, Rahimi Zahari, and Tiong Hoo Lim. Detecting diseases in chilli plants using k-means segmented support vector machine. In 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), pages 57–61. IEEE, 2019.

[6] F. Sultana, A. Sufian and P. Dutta, "Advancements in Image Classification using Convolutional Neural Network," 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, India, 2018, pp. 122-129, doi: 10.1109/ICRCICN.2018.8718718.

[7] S. Y. Chaganti, I. Nanda, K. R. Pandi, T. G. N. R. S. N. Prudhvith and N. Kumar, "Image Classification using SVM and CNN," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2020, pp. 1-5, doi: 10.1109/ICCSEA49143.2020.9132851.

[8] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 2016, pp. 2560-2567, doi: 10.1109/IJCNN.2016.7727519.

[9] Jiang Y, Chen L, Zhang H, Xiao X (2019) Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. PLoS ONE 14(3): e0214587. https://doi.org/10.1371/journal.pone.0214587

[10] Mahajan and S. Chaudhary, "Categorical Image Classification Based On Representational Deep Network (RESNET)," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 327-330, doi: 10.1109/ICECA.2019.8822133.

[11] S. B. Reddy and D. S. Juliet, "Transfer Learning with ResNet-50 for Malaria Cell-Image Classification," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0945-0949, doi: 10.1109/ICCSP.2019.8697909.

[12] N. Gouda and J. Amudha, "Skin Cancer Classification using ResNet," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2020, pp. 536-541, doi: 10.1109/ICCCA49541.2020.9250855.

[13] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu and M. Bennamoun, "Deep Learning for 3D Point Clouds: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 12, pp. 4338-4364, 1 Dec. 2021, doi: 10.1109/TPAMI.2020.3005434.
[14] Bello SA, Yu S, Wang C, Adam JM, Li J. Review: Deep Learning on 3D Point Clouds. Remote Sensing. 2020; 12(11):1729. https://doi.org/10.3390/rs12111729
[15] Achlioptas, P., Diamanti, O., Mitliagkas, I., & Guibas, L.J. (2017). Learning Representations and Generative Models for 3D Point Clouds. International Conference on Machine Learning.
[16] Liang, J. (2020, September). Image classification based on RESNET. In Journal of Physics: Conference Series (Vol. 1634, No. 1, p. 012110). IOP Publishing.
[17] Sharma, A., & Phonsa, G. (2021, April). Image classification using CNN. In Proceedings of the International Conference on Innovative Computing & Communication (ICICC).
[18] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (2013). Deep learning. Nature, 521(7385), 436-444.
[19] Zhang, Z., & Yasumoto, K. (2017). Multi-label classification using convolutional neural networks for scene understanding. IEEE Access, 5, 15601-15610.
[20] Nichol, A., Jun, H., Dhariwal, P., Mishkin, P., & Chen, M. (Year). "PointE: A System for Generating 3D Point Clouds from Complex Prompts.
[21] Hui Chen Yipeng Zuo, Yong Tong, Li Zhu College of Automation Engineering, Shanghai University of Electric Power, Shanghai 3100.27, China.
[22] Heng Zhang, Bin Zhuang, Yanli Liu School of Information Engineering East China Jiaotong University Nanchang 330013, China hbzhangheng@126.com, rzzhuangbin@163.com, hbliuyanli@126.com.