Novel YOLOv5 Model for Automatic Detection of Cowpea Leaves: Smart Agriculture

Vijaya Choudhary[1,2], *Research Scholar*, Paramita Guha[1], *Sr. Scientist,* Sunita Mishra[3], Chief Scientist
[1]*CSIR-National Physical Laboratory, New Delhi-110012, India*
[2]*Academy of Scientific & Innovative Research (AcSIR), Ghaziabad-201002, India*
[3]*CSIR-Central Scientific Instruments Organization, Chandigarh, India*
Corresponding Author: Paramita Guha (paramguha@gmail.com)

***Abstract* - Implementing artificial intelligence, specifically deep learning algorithms, to enhance agricultural productivity is a great initiative, especially in a country like India where agriculture is a crucial sector. Using TensorFlow and Keras for this purpose provides a solid foundation, given their popularity and extensive documentation. Using deep learning to identify and classify cowpea leaves can indeed streamline various agricultural processes, such as monitoring plant health, pest detection, and yield estimation. The utilization of YOLOv5, a CNN-based architecture, for the binary classification of cowpea leaves against other leaves like mangoes is a smart choice. Transfer learning can further optimize this model by leveraging pre-trained weights from similar tasks, which can significantly reduce the computational resources and time required for training. As you proceed with this experiments and model development, ensure robust data collection and preprocessing, as the quality of input data greatly influences the performance of deep learning models. Additionally, consider integrating techniques for data augmentation to further enhance the model's generalization capabilities. Continued research and development in this area can lead to significant advancements in agricultural practices, ultimately benefiting farmers and contributing to food security.**

Index Terms –Supervised learning, Precision Agriculture (PA), Cowpea leaves, classification, YOLO

I. INTRODUCTION

The detailed explanation highlights the significant shift towards the adoption of neural network models in agriculture, particularly for tasks like automatic detection of crops, fruits, and plant diseases. This shift is understandable given the advantages of neural network models, such as their ability to handle large datasets and the availability of pre-trained models, which can expedite development processes. In India, where agriculture is a primary source of employment and the sector is gradually transitioning from traditional to advanced methods, the application of computer vision and deep learning techniques holds immense potential [1]. By leveraging computer vision, farmers can benefit from more efficient and timely monitoring of crops, enabling early detection of diseases and other issues that may affect plant growth and yield [2]. The focus on cowpea in this experiment aligns with the need to enhance the productivity of pulse crops in India. Cowpea, with its high nutritional value, presents a valuable opportunity for research and improvement in crop yield[3]. Utilizing deep learning models like YOLOv5 for the detection and classification of cowpea leaves demonstrates a proactive approach towards leveraging advanced technologies to address agricultural challenges [4]. with a literature review summarizing previous work in the field, followed by a detailed discussion of the methodology and results obtained from the dataset, ensures a comprehensive presentation of research findings [5]. Such structured organization enhances the clarity and coherence of this paper, making it easier for readers to follow and understand the work. Overall, my research contributes to the growing body of knowledge in agricultural science and highlights the potential of deep learning techniques in addressing critical issues faced by farmers. By continuing to explore and refine these methodologies, researchers can further advance the field of smart agriculture, ultimately benefiting farmers and ensuring food security. Figure 1 represents entire flow of work for yolov5 model.

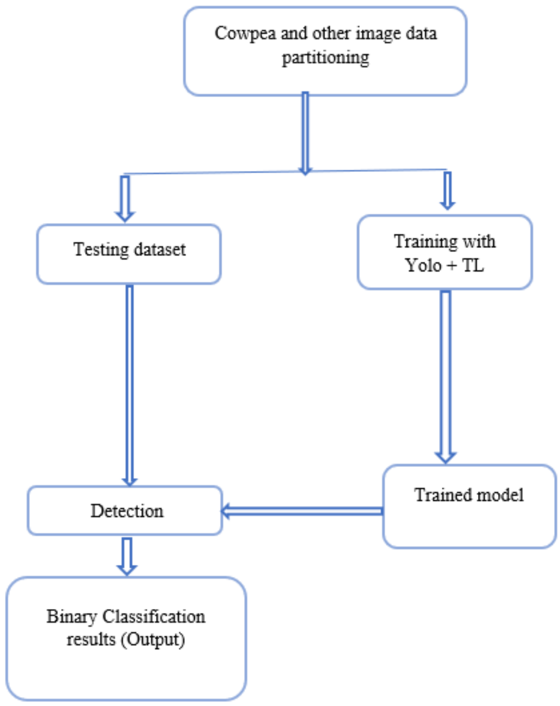

Fig. 1. Entire flow of presented work

II. LITERATURE REVIEW

This paper provides a comprehensive understanding of the evolution of image identification and classification techniques, particularly in the context of object detection [6]. Indeed, the problem of image classification has been extensively studied for years, and recent advancements in deep learning techniques have significantly improved the accuracy and efficiency of such tasks[7]. Deep learning techniques, especially Convolutional Neural Networks (CNNs), have emerged as the preferred choice for various image processing problems due to their ability to automatically learn features from data. CNNs have found applications in diverse fields[8], including natural language identification, image identification, and even handwritten digit recognition. While traditional methods like K-nearest neighbors (k-NN) and Naïve Bayes have been used for object classification in the past, they have limitations such as low efficiency, high dependence on training data[9], and sensitivity to input data. In contrast, CNNs offer superior performance by automatically extracting relevant features from images, eliminating the need for manual feature engineering. The introduction of advanced architectures like AlexNet has further improved the capabilities of CNNs,[10] enabling more accurate and efficient image recognition tasks. Additionally, the application of CNNs in tasks like automatic weed detection, as mentioned in the work by Jialin [11] showcases the versatility of these models in addressing real-world challenges in agriculture and other domains. The popularity of artificial intelligence and deep learning techniques can be attributed to the abundance of data available in various fields, enabling the training of complex models like CNNs[12]. As research in this field continues to advance, we can expect further improvements in image identification and classification techniques, ultimately leading to enhanced capabilities in areas such as object detection,[13] segmentation, and intelligent diagnosis. summary provides a valuable insight into the ongoing advancements in image classification and the increasing importance of deep learning techniques, particularly CNNs, in addressing complex problems across different domain [14][15].

III. Model for Image Detection

The description provides a clear overview of the methodology adopted in this research, particularly in the application of the YOLOv5 model for the identification of cowpea leaves. Highlighting the sophistication of the YOLOv5 model in image preprocessing and its integration with transfer learning techniques underscores the advanced nature of this approach. By leveraging a predefined convolutional neural network like YOLOv5, you benefit from its ability to handle large datasets with different object categories, ultimately aiming to achieve high accuracy in object detection tasks. The use of TensorFlow and Keras libraries in Python for importing the YOLOv5 model facilitates seamless integration into this research workflow. This choice of programming language and libraries is widely adopted in the deep learning community, ensuring compatibility and ease of implementation. this flowchart depicting the model's process, from capturing images of cowpea leaves to resizing and preprocessing them before feeding them into the YOLOv5 model, provides a visual representation of the methodology, enhancing clarity for readers. Additionally, this consideration of preprocessing techniques based on the scale of the dataset, opting for moderate preprocessing due to the relatively small dataset used in this experiments, showcases practical decision-making in research methodology. This pragmatic approach ensures efficient utilization of computational resources while still achieving satisfactory results. Overall, this methodological description lays a solid foundation for understanding the implementation of the YOLOv5 model for cowpea leaf identification, demonstrating a systematic and thoughtful approach to utilizing advanced deep learning techniques in agricultural applications. Capturing cowpea images from ICAR, Pusa Campus, New Delhi, India, using a standard DSLR camera lends credibility to your research, as these images likely represent the local variety of cowpea leaves accurately. Starting with a small dataset consisting of 100 training images and 20 testing images is a prudent approach for an initial stage of research, allowing you to assess the feasibility and validity of your model before scaling up.

Saving the images in JPG format at different resolutions ensures compatibility and ease of handling for further processing. Having two types of images, presumably cowpea leaves and another type for comparison, enables binary classification, facilitating the training and evaluation of your model. Figure 2, displaying some of the training data images, provides a visual representation of the dataset used for model training. This visualization allows researchers to gain insights into the characteristics and variability of the images, which can inform preprocessing steps and model training strategies. Overall, the presented methodology of using high-quality images from a reputable source, starting with a small dataset, and systematically evaluating your model's performance demonstrates a rigorous and systematic approach to deep learning research. This approach sets a solid foundation for further experimentation and refinement as you progress in your research and figure 3 shows the annotated image/data set for yolov5 algorithms. Entire experiments have been executed on a PC with 12 core CPU running @ 3.70 Hz with Intel processor with operating system installed Windows 10. An open source platform (Python) has been used as programming environment. Optimization algorithms play a crucial role in training neural network models, including YOLOv5 architecture, for tasks such as binary classification of images. In this paper, I applied optimization techniques to facilitate the training process. Optimization algorithms, such as Adam or RMSProp, are used to update the parameters of the neural network during training. These algorithms adjust the weights and biases of the network to minimize the loss function, improving the model's performance over time. Binary cross-entropy is a commonly used loss function for binary classification tasks. It measures the difference between the predicted probabilities and the actual binary labels, penalizing incorrect predictions and guiding the optimization process to minimize the loss. Rescaling or resizing images is essential in deep learning to ensure uniformity in input dimensions. This process reduces the number of pixels in the images, making them more manageable for the neural network. Rescaling also helps mitigate the computational complexity and training time of the model. Deep learning is a subset of machine learning that focuses on learning hierarchical representations of data through artificial neural networks. Deep learning excels in tasks involving large amounts of data, such as image classification, by automatically learning features from the data. CNNs are well-suited for image classification tasks within deep learning. They consist of convolutional layers that extract spatial features from input images, followed by pooling layers that downsample the feature maps. CNNs have been successfully applied to various applications, including object detection and face recognition. Image resolution refers to the dimensions of an image, typically represented as height (H) x width (W) x depth (D). For color images, the depth is usually three (RGB channels), while for grayscale images, it is one. High-resolution images contain more detail and information, which can improve the accuracy of the model but also increase computational complexity. This paper's utilization of optimization algorithms, loss functions, image scaling, and CNNs reflects a well-rounded approach to deep learning-based image classification. These techniques contribute to the effectiveness and efficiency of your model in identifying and classifying different types of leaves, including cowpea leaves.

Fig. 2. Images dataset for training the Model

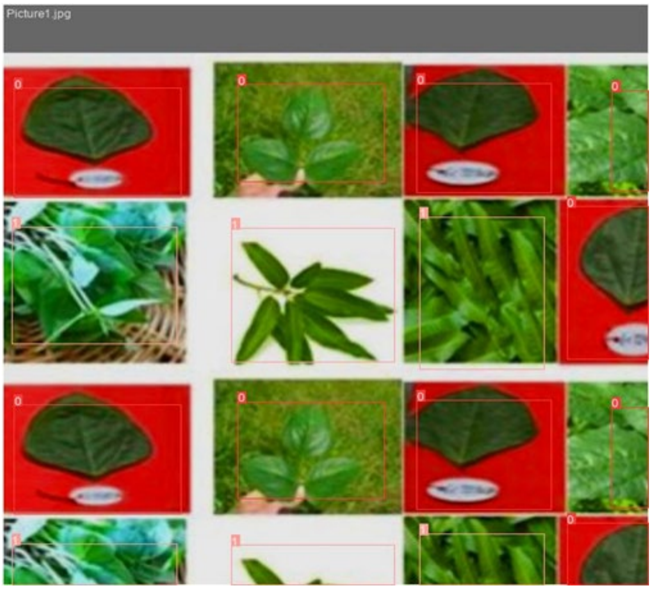

Fig. 3. Images dataset for training the Model annotated

IV. EXPERIMENTAL RESULTS

Cowpea leaf images dataset using YOLOv5 architecture over Python programming language with all Computer with adequate computational resources. The YOLOv5 algorithm, short for "You Only Look Once version 5," is an object detection algorithm that belongs to the YOLO (You Only Look Once) family of models. It is designed to efficiently detect objects in images with a single forward pass of the neural network. The algorithm takes an input image of arbitrary size. The input image is preprocessed to meet the requirements of the neural network. This typically involves resizing the image to a fixed size, normalizing pixel values, and converting it to a format suitable for feeding into the neural network. YOLOv5 utilizes a convolutional neural network (CNN) backbone to extract features from the input image. The backbone network is typically a variant of the popular architectures like ResNet, EfficientNet, or CSPNet, which are known for their ability to capture rich image features. The backbone network processes the input image and extracts features at multiple scales. These features capture both low-level details and high-level semantic information. YOLOv5's detection head consists of a series of convolutional layers responsible for predicting bounding boxes, confidence scores, and class probabilities for objects in the image. This detection head is typically appended to the backbone network. The detection head generates predictions for object detection. For each grid cell in the output feature map, the algorithm predicts bounding boxes (coordinates of the bounding box corners), confidence scores (probability that the bounding box contains an object), and class probabilities (probability distribution over different object classes). To remove duplicate detections and refine the final predictions, YOLOv5 applies non-maximum suppression. This technique selects the most confident bounding boxes while suppressing overlapping boxes with lower confidence scores. The final output of the YOLOv5 algorithm consists of a list of bounding boxes, each associated with a confidence score and a predicted class label. Optionally, post-processing steps such as filtering out low-confidence detections or applying additional constraints may be performed to improve the quality of the final detections. Overall, YOLOv5 is known for its speed and accuracy in object detection tasks, making it suitable for real-time applications where fast and precise detection of objects in images is required. The experiment focuses on distinguishing between cowpea leaves and other types of leaves (in this case, mango leaves) using a deep learning model. The model assigns a binary label of 1 for cowpea leaves and 0 for other leaves. The sigmoid function is used as the activation function in the output layer of the neural network, enabling binary classification. It maps the output values to probabilities between 0 and 1, facilitating decision-making in classification tasks. The experiment involves training the model with different numbers of epochs (i.e., training iterations). Epoch value of 3 tested to evaluate their impact on the model's accuracy. illustrate the model's performance with each optimizer. A relatively small dataset of around 60 cowpea leaf images is used for training and testing the model. The dataset is divided into 80% for training and 20% for testing. The model is trained to detect cowpea leaves using this dataset. The experimental results indicate that the model achieves an average accuracy of 80% or above when trained. Figures 5 show the training and testing accuracy curves, while Figures 7 depict the corresponding curves. The obtained results are evaluated based on the accuracy of the model in identifying cowpea leaves from the test dataset. The model's ability to correctly classify cowpea leaves is assessed using both curated images and images obtained from the internet. Overall, your explanation provides a comprehensive overview of the experimental process, results, and evaluation metrics used to assess the performance of the deep learning model in binary classification of cowpea leaves. These findings can inform future research and optimization efforts to further enhance the accuracy and efficiency of the model. Figure 6 depics results of the matrix formate using yolo algorithms. Figure 7 is also a results shown using deep learning methods.

Fig. 4. Results using YOLOv5 with annotated tags as Cowpea_Leaves and Other_Leaves

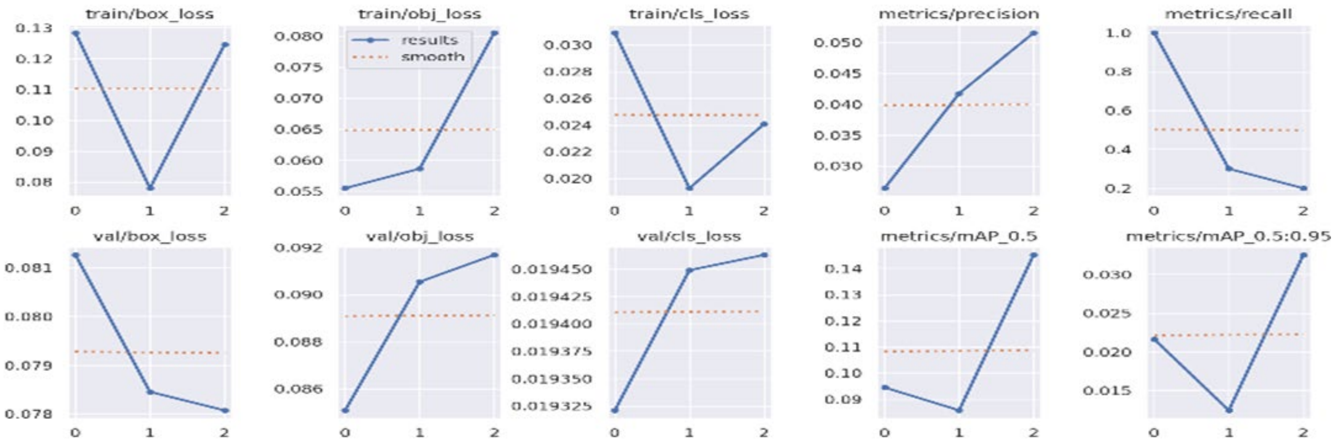

Fig. 5. Accuracy using YOLOv5 algorithms

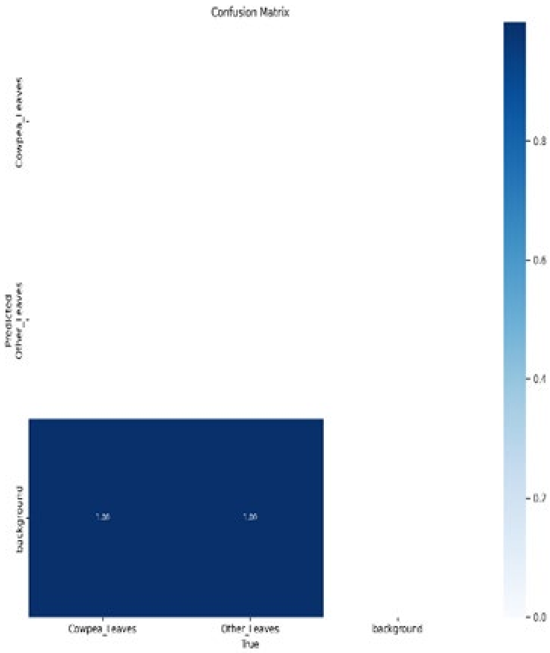

Fig. 6. Accuracy as confusion matrix using YOLOv5

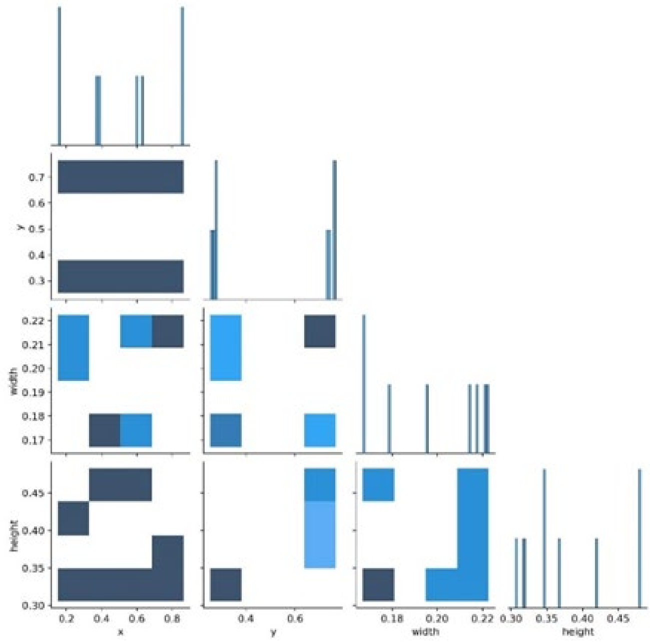

Fig. 7. Model results using YOLOv5

V. CONCLUSION

This paper's focus on utilizing a deep neural network model, specifically YOLOv5 architecture, for the identification of cowpea leaves is commendable. By leveraging advanced techniques like deep learning, you aim to enhance the accuracy of automatic detection, which can be instrumental in aiding agricultural researchers in differentiating cowpea leaves from other plants, including their wild counterparts. Highlighting the importance of expanding the database is crucial, as it not only improves the robustness and generalization capabilities of the model but also allows for better training and validation. With a larger and more diverse dataset, this proposed model can become more reliable and effective in real-world scenarios. Overall, this paper's contribution lies in bridging the gap between manual-based recognition and automated techniques, which can significantly streamline agricultural processes and research efforts. By demonstrating the efficacy of this method through experiments, you provide valuable insights into the potential applications of deep learning in the agricultural sector.

ACKNOWLEDGMENT

Acknowledging ICAR, Delhi, India for the images leaves of legume cowpea to demonstrates gratitude and recognizes the contribution of the institution to the research project. It's important to acknowledge the sources of data and support received during the course of the research, as it reflects the collaborative nature of scientific endeavors.

REFERENCES

[1] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90.

[2] Voulodimos, A., Doulamis, N., Bebis, G., & Stathaki, T. (2018). Recent Developments in Deep Learning for Engineering Applications. Computational Intelligence and Neuroscience, 2018, 1– 2

[3] Fujita, E., Kawasaki, Y., Uga, H., Kagiwada, S., & Iyatomi, H. (2016). Basic Investigation on a Robust and Practical Plant Diagnostic System. 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA).

[4] Ubbens, J., Cieslak, M., Prusinkiewicz, P., & Stavness, I. (2018). The use of plant models in deep learning: an application to leaf counting in rosette plants. Plant Methods, 14(1)

[5] JC Rana, NK Gautam, MS Gayacharan, R Yadav, K Tripathi, SK Yadav, "Genetic resources of pulse crops in India: An overview" IndianJournal of Genetics and Plant Breeding 76(4):420-436, Nov 2016

[6] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-timeobject detection with region proposal networks," in Advances in Neural Information Processing Systems 28, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 91–99.

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," inProceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.

[9] Z. Wang, P. Yi, K. Jiang, J. Jiang, Z. Han, T. Lu, and J. Ma, "Multimemory convolutional neural network for video super-resolution," IEEE Transactions on Image Processing, vol. 28, no. 5, pp. 2530–2544, 2018.

[10] J. Ma, X. Wang, and J. Jiang, "Image super-resolution via dense discriminative network," IEEE Transactions on Industrial Electronics, pp. 1–1, 08 2019, doi: 10.1109/TIE.2019.2934071.

[11] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, "Fusiongan: A generative adversarial network for infrared and visible image fusion," Information Fusion, vol. 48, pp. 11–26, 2019.

[12] J. Ma, P. Liang, W. Yu, C. Chen, X. Guo, J. Wu, and J. Jiang, "Infrared and visible image fusion via detail preserving adversarial learning," Information Fusion, vol. 54, pp. 85–98, 2020.

[13] A. Toshev and C. Szegedy, "Deeppose: Human pose estimation via deepneural networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 1653–1660.

[14] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," Mechanical Systems and Signal Processing, vol. 72, pp. 303–315, 2016.

[15] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in Proceedings of the 25th international conference on Machine learning. ACM, 2008, pp. 160–167.