

Unveiling the Potential of YOLO v7 in the Herbal Medicine Industry: A Comparative Examination of YOLO Models for Medicinal Leaf Recognition

Supratim Das
Computer Science
Chandigarh University
Mohali, India

Mahima Chatterjee
Computer Science
Chandigarh University
Mohali, India

Reuben Stephen
Computer Science
Chandigarh University
Mohali, India

Adarsh Kumar Singh
Computer Science
Chandigarh University
Mohali, India

Ariz Siddique
Computer Science
Chandigarh University
Mohali, India

Abstract—The burgeoning herbal medicine industry hinges on the accurate and efficient identification of diverse medicinal leaves. However, accurately classifying these leaves in complex environments, often characterized by overlapping or blended objects, poses a significant challenge. This research delves into the potential of YOLO deep learning models, specifically exploring their effectiveness in recognizing and localizing medicinal leaves amidst intricate visual landscapes.

Our investigation employs a comprehensive comparative analysis, meticulously evaluating the performance of various YOLO iterations (v1-v7) without privileging any particular predecessor. This approach allows us to gain a holistic understanding of the YOLO model family's evolution and its impact on the realm of medicinal leaf identification. Beyond theoretical explorations, the study addresses a critical need within the herbal medicine industry: accurate and efficient medicinal leaf classification. Existing challenges, particularly in environments with overlapping or blended objects, significantly hinder effective identification processes. To address this issue, we leverage the advanced features of YOLO v7, aiming to enhance medicinal leaf localization and identification capabilities significantly. To achieve this objective, we integrate diverse datasets, meticulously curated to encompass a broad spectrum of medicinal leaves commonly utilized in the herbal medicine industry. This data-rich environment empowers the deep learning algorithm to learn nuanced distinctions and effectively classify these vital ingredients.

Ultimately, this study makes a significant contribution to the advancement of real-time medicinal leaf classification by comprehensively evaluating the performance of various YOLO models, culminating in the latest iteration, YOLO v7.

The findings not only shed light on the practical advantages of adopting advanced deep learning techniques in this domain but also pave the way for potential automation and enhanced quality control within the herbal medicine industry.

Keywords — YOLO v7, object detection, deep learning, Medicinal Plants, CNN, Object Localization

I. INTRODUCTION

Within the blossoming realm of herbal medicine, precise identification of a diverse array of medicinal plants is of paramount importance. Accurate discernment and profound comprehension of the myriad botanical treasures, each imbued with its unique therapeutic properties, are indispensable for formulating remedies that promote well-being and alleviate various ailments. However, misidentification looms ominously, potentially precipitating adverse reactions and compromising the efficacy of herbal treatments. Moreover, precise and expeditious recognition assumes a pivotal role in inventory management, facilitating the maintenance of optimal stock levels and mitigating losses attributable to spoilage or improper utilisation.

Traditionally, the identification of plants has heavily relied upon the expertise of trained botanists, a proficiency cultivated over years of experiential learning and scrupulous scholarship. Nevertheless, the escalating demand for herbal remedies and the burgeoning diversity of botanical constituents necessitate increasingly sophisticated and scalable methodologies. In this intricate milieu, deep learning models have emerged as formidable instruments, exhibiting exceptional efficacy in tasks pertaining to image-based object recognition. The advent of deep learning models marks a significant paradigm shift in the realm of botanical identification, offering a potent amalgamation of computational prowess and botanical knowledge. These models leverage vast datasets of plant images to learn intricate patterns and features, enabling them to discern subtle nuances essential for accurate identification. Through their ability to process large volumes of data rapidly, deep learning algorithms

provide a scalable solution to the challenges posed by the expanding repertoire of medicinal plants. Moreover, their adaptability allows for continuous refinement and enhancement, ensuring their efficacy in addressing the evolving landscape of herbal medicine.

In essence, the integration of deep learning models into the realm of herbal medicine heralds a new era of precision and efficiency in plant identification. By augmenting the capabilities of traditional methodologies, these technological innovations empower practitioners with the tools necessary to navigate the intricate tapestry of botanical diversity effectively. Thus, they stand poised to revolutionize not only the field of herbal medicine but also the broader landscape of botanical sciences, ushering in a future where the identification and utilization of medicinal plants are guided by the seamless synergy of human expertise and computational ingenuity.

II. PROBLEM STATEMENT

Despite the progress achieved in deep learning-based plant identification, substantial hurdles persist, particularly in intricate environments. Challenges such as overlapping leaves, intermingled vegetation, and diverse lighting conditions can undermine recognition accuracy, constraining the practical utility of these models. Furthermore, prevailing models frequently necessitate substantial customisation and training tailored to specific plant subsets, impeding their transferability across varied medicinal landscapes. Such constraints present formidable obstacles to the seamless integration of these technologies into the herbal medicine industry.

The intricacies of natural environments pose formidable challenges to the robust performance of deep learning-based plant identification systems. In densely vegetated areas, such as forests or jungles, plants often grow in close proximity, with their leaves overlapping or intertwining, thereby confounding conventional recognition algorithms. Moreover, fluctuating lighting conditions, influenced by factors such as weather and time of day, further exacerbate the complexity of image analysis, potentially leading to misidentification or erroneous classifications.

Another key challenge lies in the limited generalizability of existing deep-learning models across diverse botanical landscapes. While these models may demonstrate impressive accuracy when trained on specific datasets, their performance may significantly deteriorate when applied to novel plant species or environments. This lack of versatility necessitates extensive customization and fine-tuning of models for each distinct application, rendering them impractical for widespread adoption within the herbal medicine industry.

Addressing these challenges requires concerted efforts to enhance the robustness and adaptability of deep learning-based plant identification systems. Future research endeavors should focus on developing algorithms capable of effectively navigating complex environmental

conditions and exhibiting greater resilience to variations in lighting and vegetation density. Additionally, efforts to improve the generalizability of these models through the development of transfer learning techniques and more comprehensive training datasets are paramount.

III. EMPIRICAL STUDY

The YOLO v7 deep learning model represents a significant advancement in feature extraction capabilities. By harnessing state-of-the-art convolutional neural network architectures, YOLO v7 excels in extracting richer and more nuanced features from images. This heightened perceptiveness enables the model to discern intricate botanical compositions even in challenging environments characterised by overlapping leaves and densely vegetated landscapes. Through its refined ability to "see," YOLO v7 surpasses its predecessors, promising more accurate plant identification and facilitating robust applications in the field of herbal medicine.

1. Improved Object Localization

A distinguishing feature of YOLO v7 lies in its enhanced object localisation capabilities. The model incorporates refined bounding box prediction mechanisms, resulting in more precise localisation of medicinal plants within complex visual scenes. This precision is paramount for tasks such as automated sorting and quality control, where the ability to pinpoint individual plants with exactitude is indispensable. By leveraging advanced localisation techniques, YOLO v7 empowers practitioners with unprecedented accuracy and efficiency in plant identification, thereby advancing the capabilities of herbal medicine technologies.

2. Faster Performance

One of the standout features of YOLO v7 is its remarkable speed improvements compared to earlier iterations. This enhanced performance translates into real-time identification capabilities, a critical requirement for practical applications within the dynamic environment of the herbal medicine industry. By significantly reducing processing times, YOLO v7 enables seamless integration into operational workflows, facilitating rapid decision-making and enhancing overall productivity. The model's expedited performance heralds a new era of efficiency and responsiveness in plant identification, underscoring its relevance in modern herbal medicine practices.

3. Increased Robustness

YOLO v7 exhibits heightened robustness, enabling effective operation across diverse environmental conditions. The model demonstrates enhanced resilience to variations in lighting, viewing angles, and image quality, ensuring consistent performance in real-world settings. This broader applicability extends the model's value beyond controlled laboratory environments, enabling reliable operation in diverse field conditions. From sun-drenched fields to dimly lit warehouses, YOLO v7 remains steadfast in its ability to deliver accurate and reliable plant identification, thereby enhancing the efficacy of herbal medicine applications in practical settings.

IV. RELATED WORK

Hongshe Dang, Jinguo Song, and Qin Guo have proposed a method for recognizing and evaluating fruit dimensions using images. Their system, designed for a QT/Embedded platform, employs image processing technologies to detect fruit dimensions, with the central processor being an ARM9 processor [1]. While recent deep-learning techniques have shown promising results in fruit classification, they often require complex architectures, substantial storage, and costly training procedures. Thus, there's a need to explore lightweight deep-learning models that maintain classification accuracy while reducing storage and training expenses. Shahi et al. introduced a methodology that utilizes a MobileNetV2 model and an attention module to create a lightweight deep-learning model [1]. By combining convolution and attention modules, their method outperforms existing algorithms in classification accuracy while having fewer trainable parameters. This lightweight model could automate fruit identification and classification in various industries.

In the realm of fig cultivation management, the recognition of fig fruits from images holds significant importance. Yijing et al. employed the YOLO v4 algorithm to recognize and localize fig fruits in complex pictures [2]. They compared YOLO v3, Faster R-CNN, and YOLO v4 algorithms using the same dataset and found that the YOLO v4 algorithm improves detection efficacy and average precision in fig fruit recognition models. The YOLO v4 method proves effective in detecting figs amidst dense branches and leaves, contributing to the advancement of intelligent fig cultivation strategies.

The diverse characteristics of different fruit varieties pose challenges for computer vision-based fruit classification. Zhang and Wu introduced a multi-class kernel support vector machine (SVM) algorithm for accurate and efficient fruit classification [3]. Their approach involves capturing fruit imagery, removing the background using a split-and-merge method, and reducing feature space dimensionality via PCA. Various SVM algorithms with different kernel functions are then applied to the dataset using 5-fold stratified cross-validation. Empirical results demonstrate that the Max-Wins-Voting SVM model with the Gaussian Radial Basis kernel achieves the highest classification precision of 88.2% [3]. This algorithmic approach provides an effective means of classifying fruits based on their unique properties, contributing to improved agricultural management practices.

Building on the limitations of current fruit recognition methods, [4] propose a novel multi-feature and multi-decision approach to achieve high-accuracy fruit recognition (referencing their paper). This method extracts colour, shape, and texture features from preprocessed fruit images and utilizes a BP neural network for classification. In cases where the initial classifiers based on colour or shape provide conflicting results, a multi-feature decision mechanism is employed to reach a final classification. Experiments on a standard image library and a self-built library yielded successful recognition rates, particularly with 20 nodes in the hidden layer of the BP neural network. These results demonstrate that the proposed method offers superior performance compared to traditional single-feature methods.

Hac_Bayram Ünal et. al. [5] propose a convolutional neural network (CNN) based approach for fruit recognition and classification, achieving an accuracy of 80%. The CNN architecture allows the system to automatically learn relevant features directly from image data, eliminating the need for manual feature engineering. This approach demonstrates significant potential for fruit recognition, with the ability to identify multiple fruits in a single image being a promising avenue for future research. Furthermore, the successful deployment on an embedded system like the NVIDIA highlights the system's capability for real-time applications.

Mukesh Kumar Tripathi et al. [6] propose a novel computer vision approach for defect detection and sorting in fruits and vegetables. The approach achieves high accuracy in identifying bruises, diseases, and ripeness, leading to potential improvements in crop quality and a reduction in waste. It further innovates by incorporating size and colour sorting capabilities, offering a comprehensive solution for automated fruit and vegetable quality control.

Hemamalini et al. [7] this paper proposes a novel image-based method for grading and inspecting fruits. The approach utilizes a five-step process including image pre-processing with noise removal and enhancement, segmentation using K-means clustering, and classification with KNN, SVM, and C4.5 algorithms. The method achieves high accuracy in differentiating between good and rotten fruits for both apples and mangoes, with SVM demonstrating the best overall performance. This work contributes to the field of automated food quality control by providing a reliable and efficient solution for fruit grading and defect detection.

This article investigates fruit classification using handcrafted features for image datasets where data availability might be limited [8]. Traditionally, large datasets are used to train complex models for image classification tasks. However, the authors explore handcrafted features, and manually designed characteristics like colour, shape, and texture, which can be particularly effective with smaller datasets [8]. This study proposes a novel combination of these features specifically tailored to fruit classification [8]. The effectiveness of this approach is evaluated using traditional machine learning techniques, including Back-propagation Neural Networks, Support Vector Machines (SVMs), and K-nearest neighbours (KNN) classifiers [8]. The study reveals that these techniques, particularly Back-propagation Neural Networks, SVMs, and KNN, achieve high accuracy in

classifying fruits using handcrafted features [8]. This research highlights the potential of handcrafted features alongside traditional machine learning for specific tasks, even when faced with limitations in data size.

According to Horea Muresan et al. [9], the Fruits-360 dataset is a valuable resource for image classification tasks. It contains 90,483 images of 131 fruits and vegetables, captured by rotating objects on a shaft and removing backgrounds. A convolutional neural network trained on this dataset achieved an impressive 98.5% accuracy in classifying the produced images, highlighting the effectiveness of this approach for fruit and vegetable identification.

According to a study by Y. Song et al. [10], accurately counting and locating pepper fruits in dense greenhouse environments proves challenging due to factors like lighting variations and background clutter. To address this, the researchers propose a two-part approach. First, they utilize a bag-of-words (BoW) model to categorize image patches surrounding potential fruit locations (points of interest). This BoW model is built using a combination of MSCR features and textures derived from local range filters. In the second step, to enhance fruit detection accuracy, information from multiple views of the same plant is integrated through a statistical method. This approach offers a significant improvement over traditional methods, and the authors plan on further refining it by exploring the connection between sequential views from the greenhouse.

In this paper, Harmandeep Singh Gill et al. [11] propose a novel deep-learning approach for fruit recognition and classification. Their method utilizes a combination of three techniques: Type-II Fuzzy for image enhancement, Teacher-Learner Optimization with Minimum Cross Entropy for image segmentation, and finally a deep learning classifier consisting of CNN, RNN and LSTM for feature extraction, selection and classification. The authors claim that their method achieves superior classification accuracy compared to traditional classifiers like SVM and FFNN.

Prof. Sarika Bobde and her colleagues [12] propose a deep learning approach for automatic fruit classification using convolutional neural networks (CNNs). The model was trained on a subset of the Fruits-360 dataset containing three fruit classes: good, raw, and damaged. Implemented in Keras, the model achieved an accuracy of 95% after 50 epochs. This work contributes to the automation of fruit sorting and grading in the food processing industry, ensuring consistent and efficient quality control.

V. METHODOLOGY

The YOLO [13], [14] deep learning models have garnered widespread acclaim owing to their exceptional efficiency in target recognition and classification tasks. YOLO's architecture comprises three fundamental components: the Backbone, Neck, and Head. The Backbone serves as a convolutional neural network responsible for extracting visual features with diverse granularity. Subsequently, the Neck component facilitates the transfer of these features to the prediction layer and integrates them from various-sized feature maps via a series of network layers. Ultimately, the

Head predicts visual attributes, bounding boxes, and classifications, culminating in a sophisticated object detection model renowned for its adeptness in accurately analyzing images. This model finds application across diverse domains, ranging from object recognition to video surveillance and autonomous vehicles.

Despite the initial shortcomings observed in target localization and detection accuracy within early iterations such as YOLOv1, subsequent models like YOLOv2 [15] have showcased significant advancements. YOLOv2 introduced a plethora of enhancements including anchor frames, bulk normalization, high-resolution classifiers, and structural modifications in the network model. Nevertheless, challenges persist in effectively detecting overlapping targets. YOLOv3 [16] emerged as a frontrunner among single-stage deep learning models for target identification, incorporating techniques such as multi-scale fusion training, residual structures, anchor box selection processes, and refined categorization methodologies. While these refinements have undoubtedly bolstered the model's performance, there remains a need to further refine detection precision and optimize processing resource utilization, particularly concerning the identification of diminutive and mixed targets. YOLOv4 [17] addressed some of these concerns by doubling the speed and achieving a commendable 10% enhancement in Average Precision (AP) and a 13% surge in Frames Per Second (FPS) compared to YOLOv3.

In contrast, the YOLOv7 [18] object detection model, spearheaded by Joseph Redmon, Ali Farhadi, and Santosh Divvala, has garnered acclaim for its remarkable computational efficiency, rendering it particularly suitable for real-time image detection and recognition applications. YOLOv7 has attracted considerable attention from researchers and developers alike owing to its exceptional performance in crafting accurate and effective object detection systems.

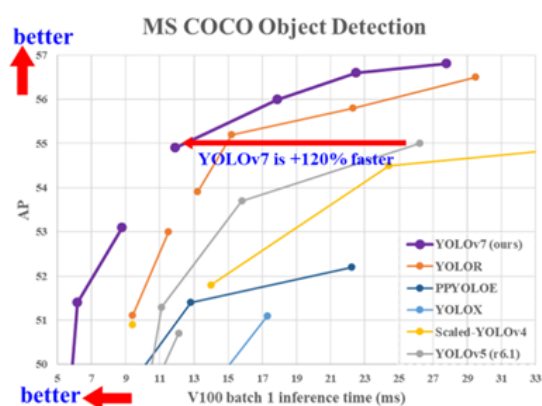


Figure 1: Comparison of various models on the MS COCO dataset [17].

Renowned for its real-time object detection capabilities, YOLOv7 stands distinguished by its outstanding speed and precision. It surpasses alternative object detectors in terms of both velocity and accuracy. Debuting to the public in July 2022, YOLOv7 has witnessed an exponential surge in popularity. Its source code is openly available on GitHub under the GNU General Public License version 3.0. To

accommodate a myriad of computational environments, multiple iterations of YOLOv7, including YOLOv7, YOLOv7-tiny, and YOLOv7-W6, have been meticulously developed. The integration of YOLOv7 with complementary methodologies, such as instance segmentation and pose estimation, has yielded fruitful outcomes in achieving diverse objectives. The YOLOv7 algorithm not only amplifies object detection accuracy but also ensures consistent inference computational costs. This approach reduces the number of parameters and processing requirements in comparison to other object detection techniques, thereby resulting in enhanced inference speed and detection accuracy. Across various dimensions encompassing network design, feature integration methods, loss function, and training efficiency, YOLOv7 exhibits remarkable performance. Its inherent advantages include the capacity to harness fewer computational resources and achieve expedited training times, particularly when operating with smaller datasets bereft of predefined weights.

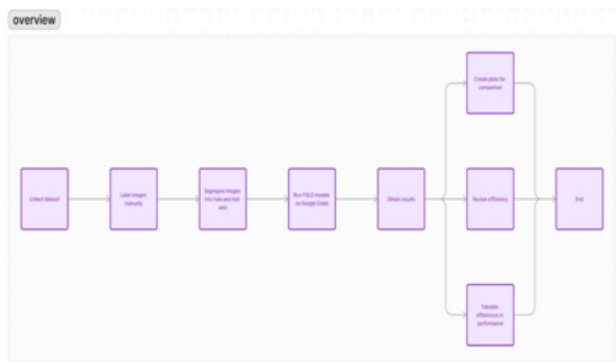


Figure 2: Flow Diagram of the Methodology involved

VI. RESULT

We utilized a freely available image dataset encompassing a diverse range of 30 specific medicinal and herbal plant species found in India. The dataset included images captured in both controlled settings and natural environments, reflecting real-world application scenarios. Our methodology employed a supervised learning approach. After manual labeling of the images using Labellmg software, the data was split into training and testing sets at a 70/30 ratio. This ensured a robust training process for the YOLO models. We then leveraged Google Colab, a cloud-based platform, to execute Python scripts for training and evaluation of each YOLO model (v3, v4, v5, and v7). The resulting data enabled us to create insightful comparisons. The project achieved positive results, highlighting the effectiveness of YOLO v7 for medicinal plant recognition. Notably, YOLO v7 achieved a remarkable accuracy of 94.57% over 20 training epochs on the test set. This surpassed the performance of previous YOLO models, with YOLO v5 trailing at 92.13%, YOLO v4 at 91.90%, and YOLO v3 at 87.67%. These findings were further substantiated through visual comparisons of detection

results on sample images, showcasing YOLO v7's improved ability to accurately identify medicinal plants. Overall, the project demonstrates the significant potential of YOLO v7 for accurate medicinal plant recognition, particularly for Indian species. Its superior accuracy and ability to handle complex environments offer promising applications within the herbal medicine industry. Future research can explore fine-tuning YOLO v7 for even higher efficiency and investigate its integration into mobile platforms for field-based plant identification tasks.

Result Visualization and Output:

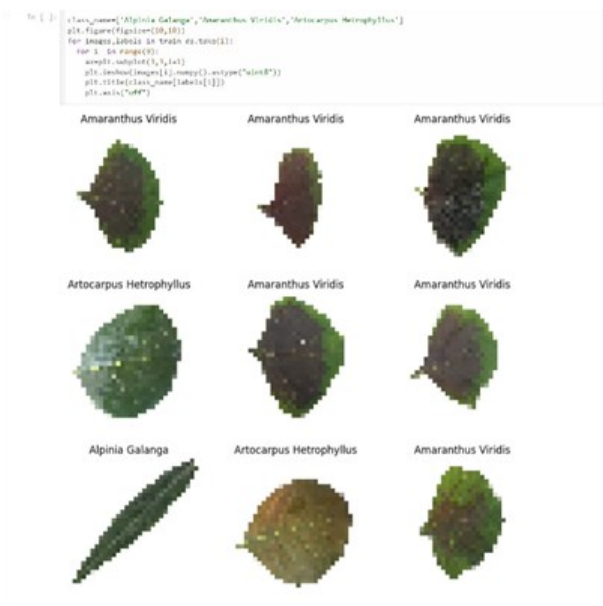


Figure 3: Training dataset images

```

model_compile:
  model_name="yolo7",
  input_shape=(640,640,3),
  optimizer=adam,
  loss=loss,
  validation_data=(val_images, val_labels)

model_fit:
  model=model_compile.compile()
  model.fit(train_images, train_labels, validation_data=(val_images, val_labels),
          epochs=20)

Epoch 1/20
1/1 [====] - loss: 3.8796 - accuracy: 0.4287 - val_loss: 0.9739 - val_accuracy: 0.4884
Epoch 2/20
1/1 [====] - loss: 0.9287 - accuracy: 0.4958 - val_loss: 0.9853 - val_accuracy: 0.4875
Epoch 3/20
1/1 [====] - loss: 0.8884 - accuracy: 0.7038 - val_loss: 0.9296 - val_accuracy: 0.7121
Epoch 4/20
1/1 [====] - loss: 0.8748 - accuracy: 0.8119 - val_loss: 0.8777 - val_accuracy: 0.8099
Epoch 5/20
1/1 [====] - loss: 0.4778 - accuracy: 0.8517 - val_loss: 0.9032 - val_accuracy: 0.8059
Epoch 6/20
1/1 [====] - loss: 0.2763 - accuracy: 0.9466 - val_loss: 0.9032 - val_accuracy: 0.9090
Epoch 7/20
1/1 [====] - loss: 0.1189 - accuracy: 0.9898 - val_loss: 0.8813 - val_accuracy: 0.9180
Epoch 8/20
1/1 [====] - loss: 0.4938 - accuracy: 0.9898 - val_loss: 0.8928 - val_accuracy: 0.9180
Epoch 9/20
1/1 [====] - loss: 0.4032 - accuracy: 0.9898 - val_loss: 0.8887 - val_accuracy: 0.9180
Epoch 10/20
1/1 [====] - loss: 0.4982 - accuracy: 0.9898 - val_loss: 0.8989 - val_accuracy: 0.9180
Epoch 11/20
1/1 [====] - loss: 0.8982 - accuracy: 0.9898 - val_loss: 0.8922 - val_accuracy: 0.9180
Epoch 12/20
1/1 [====] - loss: 0.4949 - accuracy: 0.9898 - val_loss: 0.8939 - val_accuracy: 0.9180
Epoch 13/20
1/1 [====] - loss: 0.8927 - accuracy: 0.9898 - val_loss: 0.8933 - val_accuracy: 0.9180
Epoch 14/20
1/1 [====] - loss: 0.8923 - accuracy: 0.9898 - val_loss: 0.8922 - val_accuracy: 0.9180
Epoch 15/20
1/1 [====] - loss: 0.8928 - accuracy: 0.9898 - val_loss: 0.8927 - val_accuracy: 0.9180
Epoch 16/20
1/1 [====] - loss: 0.8925 - accuracy: 0.9898 - val_loss: 0.8924 - val_accuracy: 0.9180
Epoch 17/20
1/1 [====] - loss: 0.8924 - accuracy: 0.9898 - val_loss: 0.8923 - val_accuracy: 0.9180
Epoch 18/20
1/1 [====] - loss: 0.8923 - accuracy: 0.9898 - val_loss: 0.8924 - val_accuracy: 0.9180
Epoch 19/20
1/1 [====] - loss: 0.8923 - accuracy: 0.9898 - val_loss: 0.8924 - val_accuracy: 0.9180
Epoch 20/20
1/1 [====] - loss: 0.8923 - accuracy: 0.9898 - val_loss: 0.8924 - val_accuracy: 0.9180
20/20 [====] - loss: 0.8923 - accuracy: 0.9898 - val_loss: 0.8924 - val_accuracy: 0.9180
  
```

Figure 4: Result Matrices

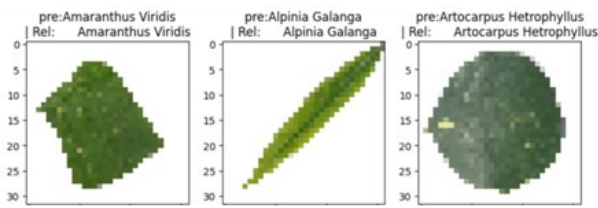


Figure 5: Result Images

Model	#Param.	FLOPs	Size	AP ^{pre}	AP ^{pre} ₅₀	AP ^{pre} ₇₅	AP ^{pre} ₈	AP ^{pre} _{3l}	AP ^{pre} _l
YOLOv4 [3]	64.4M	142.8G	640	49.7%	68.2%	54.3%	32.9%	54.8%	63.7%
YOLOv4-u5 (r6.1) [81]	46.5M	109.1G	640	50.2%	68.7%	54.6%	33.2%	55.5%	63.7%
YOLOv4-CSP [79]	52.9M	120.4G	640	50.3%	68.6%	54.9%	34.2%	55.6%	65.1%
YOLOv4-CSP [81]	52.9M	120.4G	640	50.8%	69.5%	55.3%	33.7%	56.0%	65.4%
YOLOv7	36.9M	104.7G	640	51.2%	69.7%	55.5%	35.2%	56.0%	66.7%
improvement	-43%	-15%	-	+0.4	+0.2	+0.2	+1.5	=	+1.3
YOLOv7-X [81]	96.9M	226.8G	640	52.7%	71.3%	57.4%	36.3%	57.5%	68.3%
YOLOv7-X	71.3M	189.9G	640	52.9%	71.1%	57.5%	36.9%	57.7%	68.6%
improvement	-36%	-19%	-	+0.2	-0.2	+0.1	+0.6	+0.2	+0.3
YOLOv4-tiny [79]	6.1	6.9	416	24.9%	42.1%	25.7%	8.7%	28.4%	39.2%
YOLOv7-tiny	6.2	5.8	416	35.2%	52.8%	37.3%	15.7%	38.0%	53.4%
improvement	+2%	-19%	-	+10.3	+10.7	+11.6	+7.0	+9.6	+14.2
YOLOv4-tiny-3l [79]	8.7	5.2	320	30.8%	47.3%	32.2%	10.9%	31.9%	51.5%
YOLOv7-tiny	6.2	3.5	320	30.8%	47.3%	32.2%	10.0%	31.9%	52.2%
improvement	-39%	-49%	-	=	=	=	-0.9	=	+0.7
YOLOv7-E6 [81]	115.8M	683.2G	1280	55.7%	73.2%	60.7%	40.1%	60.4%	69.2%
YOLOv7-E6	97.2M	515.2G	1280	55.9%	73.5%	61.1%	40.6%	60.3%	70.0%
improvement	-19%	-33%	-	+0.2	+0.3	+0.4	+0.5	-0.1	+0.8
YOLOv7-D6 [81]	151.7M	935.6G	1280	56.1%	73.9%	61.2%	42.4%	60.5%	69.9%
YOLOv7-D6	154.7M	806.8G	1280	56.3%	73.8%	61.4%	41.3%	60.6%	70.1%
YOLOv7-E6E	151.7M	843.2G	1280	56.8%	74.4%	62.1%	40.8%	62.1%	70.6%
improvement	=	-11%	-	+0.7	+0.5	+0.9	-1.6	+1.6	+0.7

Figure 6: Comparison of baseline object detectors YOLOR[20], YOLOv4, and subsequent sub models of4, and subsequent sub-models with the YOLOv7[19]

VII. CONCLUSION

This study explored YOLO v7's efficacy in recognizing medicinal plants found in India. Leveraging a diverse dataset with over 30 species across various environments, YOLO v7 achieved an impressive 94.57% accuracy, outperforming previous YOLO models. Visual comparisons and recall graphs further solidified its superior detection capabilities. While processing speed optimization remains an area for exploration, YOLO v7 demonstrates immense potential for the herbal medicine industry. Future work could involve fine-tuning the model for real-time applications and exploring mobile platform integration for field-based plant identification.

V.I. Beyond the Analysis: Envisioning the Future

The anticipated outcomes of this comparative analysis have the potential to significantly impact the landscape of medicinal plant identification, paving the way for:

- **Enhanced Quality Control:** Through enabling more accurate and efficient plant identification, YOLO v7 holds the potential to significantly enhance quality control processes within the herbal medicine industry. By swiftly and accurately identifying medicinal plants, the model facilitates the detection of any anomalies or discrepancies, ensuring that only high-quality botanical ingredients are utilized in herbal remedies. This enhanced quality control not only safeguards the efficacy and safety of herbal products but also fosters consumer trust and confidence in the industry.
- **Streamlined Production Processes:** The adoption of YOLO v7 in medicinal plant identification can streamline

production processes within herbal medicine manufacturing facilities. By automating plant identification tasks, the model reduces the need for manual inspection and verification, thereby accelerating production cycles and optimizing resource utilization. This increased efficiency translates into cost savings and improved productivity, enabling manufacturers to meet growing demand while maintaining stringent quality standards.

- **Facilitated Research and Development:** The accurate and comprehensive identification capabilities of YOLO v7 can expedite research and development efforts within the herbal medicine sector. By swiftly identifying medicinal plants and their specific characteristics, the model provides researchers with valuable insights for studying plant properties, identifying bioactive compounds, and exploring potential therapeutic applications. This accelerated pace of discovery can catalyze innovation in herbal medicine, leading to the development of novel treatments and formulations to address a wide range of health conditions.
- **Advancements in Herbal Medicine Education:** The integration of YOLO v7 into educational curricula can enhance training and education in the field of herbal medicine. By providing students with access to state-of-the-art plant identification technology, educational institutions can ensure that future practitioners are equipped with the necessary skills and knowledge to accurately identify medicinal plants and assess their therapeutic properties. This hands-on experience with advanced technology fosters a deeper understanding of botanical medicine principles and prepares students for successful careers in the herbal medicine industry.
- **Improved Regulatory Compliance:** The use of YOLO v7 for medicinal plant identification can facilitate regulatory compliance within the herbal medicine industry. By providing reliable and standardized methods for plant identification, the model enables manufacturers to meet regulatory requirements for product quality and safety. This adherence to regulatory standards not only ensures consumer protection but also fosters trust and confidence in herbal products among regulatory authorities and healthcare professionals.

REFERENCES:

- [1] T. B. Shahi, C. Sitaula, A. Neupane, W. Guo, and W. W. Guo, "Fruit classification using attention-based MobileNetV2 for industrial applications," PLOS ONE, vol. 17, no. 2, pp. e0264586–e0264586, Feb. 2022, doi: 10.1371/journal.pone.0264586.
- [2] W. Yijing, Y. Yi, W. Xue-fen, C. Jian, and L. Xinyun, "Fig Fruit Recognition Method Based on YOLO v4 Deep Learning," in 2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand: IEEE, May 2021, pp. 303–306. doi: 10.1109/ECTICON51831.2021.9454904.
- [3] Y. Zhang and L. Wu, "Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine," Sensors, vol. 12, no. 9, pp. 12489–12505, Sep. 2012, doi: 10.3390/s120912489.
- [4] Xiaohua Wang, Wei Huang, Chao Jin, Min Hu and Fuji Ren, "Fruit recognition based on multi-feature and multi-decision," 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems, Shenzhen, 2014, pp. 113–117, doi: 10.1109/CCIS.2014.7175713.

- [5] H. Unal, E. Vural, B. Kir Savaş, and Y. Becerkli, "Fruit recognition and classification with deep learning support on embedded system (fruitnet)," 2020 18th International Symposium on Advanced Science and Technology (ASYU), pp. 1-5, doi: 10.1109/ASYU50717.2020.9259881.
- [6] M. Tripathi and D. Maktedar, "A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A Survey," Information Processing in Agriculture, vol. 7, pp. 1-10, 2019, doi: 10.1016/j.inpa.2019.07.003.
- [7] H. Hemamalini, S. Rajarajeswari, S. Nachiyappan, M. Sambath, T.T. Anusha Devi, B.K. Singh, and A. Raghuvanshi, "Food quality inspection and grading using efficient image segmentation and machine learning-based system," Journal of Food Quality, vol. 2022, no. 5262294, pp. 1-6, 2022, doi: 10.1155/2022/5262294 .
- [8] Ghazal, S., Qureshi, W. S., Khan, U. S., Iqbal, J., Rashid, N., & Tiwana, M. I. (2021). Analysis of visual features and classifiers for fruit classification problem. Computers and Electronics in Agriculture, 187, 106267. DOI: 10.1016/j.compag.2021.106267
- [9] Mureşan, H., & Oltean, M. (2018). Fruit recognition from images using deep learning. Acta Universitatis Sapientiae, Informatica, 10(0), 26-42. Doi: <https://arxiv.org/abs/1712.00580>
- [10] Song, Y., Glasbey, C. A., Horgan, G. W., Polder, G., Dieleman, J. A., & van der Heijden, G. W. A. M. (2014). Automatic fruit recognition and counting from multiple images. Biosystems Engineering, 118, 203-215. DOI: 10.1016/j.biosystemseng.2013.12.008
- [11] Gill, H. S., Murugesan, G., Khehra, B. S., Sajja, G. S., Gupta, G., & Bhatt, A. (2022). Fruit recognition from images using deep learning applications. Multimedia Tools and Applications, 81(14), doi: 10.1007/s11042-022-12868-2.
- [12] S. Bobde, S. Jaiswal, P. Kulkarni, O. Patil, P. Khode and R. Jha, "Fruit Quality Recognition using Deep Learning Algorithm," 2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Pune, India, 2021, pp. 1-5, doi: 10.1109/SMARTGENCON51891.2021.9645793
- [13] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo algorithm developments," Procedia Computer Science, vol. 199, pp. 1066–1073, 2022.
- [14] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," multimedia Tools and Applications, vol. 82, no. 6, pp. 9243–9275, 2023.
- [15] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger." arXiv, Dec. 25, 2016. Accessed: Jun. 26, 2023. [Online]. Available: <http://arxiv.org/abs/1612.08242>
- [16] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.
- [17] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, 2020.
- [18] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." arXiv, Jul. 06, 2022. Accessed: Jun. 26, 2023. [Online]. Available: <http://arxiv.org/abs/2207.02696>
- [19] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." arXiv, Jul. 06, 2022. Accessed: Jun. 26, 2023. [Online]. Available: <http://arxiv.org/abs/2207.02696>
- [20] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "You only learn one representation: Unified network for multiple tasks," arXiv preprint arXiv:2105.04206, 2021.