Sugarcane stem node detection using computer vision and convolutional transfer learning

Arun Kumar
Govind Ballabh Pant University of Agriculture and Technology
Abhishek Kumar Saini
Govind Ballabh Pant University of Agriculture and Technology
Alaknanda Ashok
Govind Ballabh Pant University of Agriculture and Technology
Puneet Mishra (Puneet.mishra@wur.nl)
Wageningen University & Research

Research Article

Keywords: Computer vision, digital agriculture, detection, deep learning

Posted Date: September 12th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3326826/v1

Abstract

Quantity of seed needed for sugarcane planting can be reduced by employing the single bud planting method. This approach involves utilizing the stem node or bud extracted from harvested sugarcane for replanting. Currently, this process is manually done using traditional planting techniques. To automate node/bud planting, this study developed a computer vision system based on artificial intelligence. The purpose of this system is to automate the recognition of the stem node, where the sugarcane buds are naturally located. By automating the recognition process, this system can facilitate the automation of node cutting equipment and prevent bud damage during the cutting phase. In this study, a transfer learning-based approach was employed, harnessing the capabilities of a pre-trained convolutional neural network to adapt to the specific application of stem node detection. Transfer learning was chosen to avoid the need for developing and training a new model from scratch, thereby saving time in model development, and enhancing accuracy. Transfer learning allows us to leverage the knowledge acquired from a previously trained machine learning model and apply it to a different but related problem. The goal is to exploit the learned knowledge from one task to improve generalization in another. Specifically, the weights learned by a network during "task A" are transferred to a new "task B." Four different transfer learning models were trained and tested in this work: VGG16, ResNet-50, EfficientNetb7, and MobileNet. Practical tests demonstrated that VGG16 exhibited the best performance in stem node detection accuracy, achieving approximately 98%, with an average image processing time of 0.182 seconds per detection. The developed model and methodology hold broad applicability for automating stem node/bud-based sugarcane replanting.

Introduction

Sugarcane holds the distinction of being the world's largest crop in terms of production quantity. India plays a crucial role in the sugar industry and stands as the second-largest producer of sugarcane globally [1]. In the cultivation of sugarcane, one of the prevailing practices involves planting billets that contain buds and rings. However, it has been observed that planting sugarcane from a single bud can enhance output while enabling farmers to sell additional internodes to sugar mills for resource generation. Buds, which are situated on each node of the sugarcane, represent the vegetative portion from which the plant grows, typically arranged in an alternate pattern [2]. Direct node sowing or transplanting sugarcane buds can significantly reduce the weight-based seed requirement of sugarcane by approximately 85%, with minimal risk of bud infection [3].

Currently, there exist several manually and motorized sugarcane node cutting and bud chipping equipment in the market. While such machines are well-suited for small-scale farmers, they encounter challenges when it comes to throughput during node cutting for larger farms. Furthermore, manual cutting of sugarcane into billets results in low-quality seeds and reduced productivity during cultivation. Hence, it becomes imperative to automate the high throughput cutting of sugarcane nodes, which would save time and reduce labor requirements. Although cutting and chipping mechanization already exist, the primary challenge lies in accurately detecting the sugarcane stem node to enable machines to perform precision cutting.

The sugarcane node has a distinct appearance, including a specific color and shape, making it easy to detect with the naked eye. Therefore, an ideal solution to detect and locate the bud is to use RGB color cameras or computer vision [4]. Computer vision can be combined with the latest deep learning techniques for object detection and location, where the object is the sugarcane buds in this case [5]. Deep learning models are based on neural networks, which are inspired by the human brain's functions [6]. Neural networks consist of different types of neurons stacked on top of one another to create layers of neurons [7]. A deep neural network is created by linking more of these layers together. The ANN, CNN, and GAN are different types of deep neural networks. The CNN is primarily used for imaging data, allowing it to capture both low and high-level features in images [8]. The CNN-based deep learning models use all the variables, including color, shape, node section diameter, and other features, independently. Therefore, a CNN-based model provides improved recognition accuracy and can be effectively generalized for various sugarcane circumstances and types. Several studies have reported the use of CNN in crop disease imaging and identification [9–11], but very few have reported the use of computer vision CNN-based methods for sugarcane stem node detection. [12] achieved 91.55% precision with an SVM-based classification and recognition method, while [13] achieved 98.31% detection accuracy with hyperspectral imaging technology. [14] used an improved YOLO v3 and achieved 90.38% accuracy. However, most of these studies were conducted in a laboratory setting. Also, most of the works were conducted using techniques such as classification and recognition methods based on SVM, YOLO v3, Sobel edge detection method, etc. but training the models from scratch.

Sugarcane nodes can be identified by counting the number of pixels in an image using the HSV method, and the Sobel edge detector method is 93% accurate [15, 16]. Both methods are effective, but they may underperform or fail to detect if used with different types of sugarcane or if the bud-bearing section (node) contains an undetached leaf. The color of the node was used to detect it, and the recognition accuracy was 83.33%. Displacement sensors based on strain gauges, acceleration sensors, and piezoelectric film sensors have also been used [16–18]. However, these methods may not generalize with changes in sugarcane varieties, and identification may be impossible if some leaf sheaths remain attached to the node part. In the presented work, transfer learning was used to adapt an already trained CNN model to detect sugarcane nodes. Transfer learning [19, 20] has several benefits, including the ability to train the model with a smaller number of samples, not needing to create a model from scratch, and the model having already seen large variations as it has been pre-trained on diverse samples. The present study hypothesizes that a CNN based on transfer learning with minimal data and high recognition accuracy can allow for sugarcane node detection with a small training set. It is expected that the model can effectively generalize to various sugarcane circumstances and types and can carry out detection at any desired location.

Materials and methods

3.1. Data collection and pre-processing

Sugarcane stalks of 10 different varieties having varying characteristics like stem color, girth diameter etc., were selected from the field of the Crop Research Center at Govind Ballabh Pant University of Agriculture & Technology in Pantnagar, India. Before imaging, the samples were pre-cleaned by removing the leaf sheaths. However, the surface of the sugarcane covered in a waxy black material was not cleansed, as this is consistent with how farmers typically handle sugarcane before cutting it for planting in the field. A sugarcane stalk consists of two major parts: nodes and internodes, with the length of the internodes varying. To capture photographs of the sugarcane, a 64-megapixel camera equipped with a Sony IMX 355 sensor and an f/2.2 lens was used. The camera was from a Poco X4 Pro 5G smartphone. A physical setup was created to hold the camera 11 cm high and horizontal to the surface to take the pictures. By rotating the sugarcane at various angles, 689 photos of nodes of the sugarcane stem were captured. Similarly, 457 pictures of the internodes (the area between two nodes) were taken. Each variety (as we had multiple varieties) almost have equal number of images in the dataset. The setting was well-lit with a 7-watt LED bulb placed 20 cm from the data collection location. Sugarcanes were placed on white sheet of paper for augmenting white background in the images and to reduce background noise. For real-time model testing, the camera was connected to an Acer laptop (model – Aspire) with an Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz processor and an integrated 2 GB Intel HDD graphics card. The Iriun application was used in both the cellphone and laptop to use the cellphone camera as an external webcam. A sample image of a node (left) and an image of an internode (right) section before cropping is shown in Fig. 1. The imaging was done in a lab setting.

All data pre-processing was performed in Python program. In data pre-processing, firstly, the node portion and the internode portion were cropped out and the obtained data after cropping was reshaped to size 120 x 60 x 3 pixels. Main reasons for doing this preprocessing were, a.) to reduce the background noise, b.) to reduce the image size which will ultimately help in reducing computation and c.) to narrow down the region of interest for detection. From images of initial size: 640 × 480 x 3 pixels, the final dataset had images of size 120 x 60 x 3 pixels. Then total of 100 images, 50 each from nodes and internodes were separated from the training dataset for evaluating the models post training. After this, images were assigned to one of two groups: a.) nodes, and b) no nodes. We did this because we were attempting to categorize images and needed to approximate the location of a node. Following that, all the images were converted into arrays using several Python libraries, including NumPy, OpenCV, Keras. Labels were one-hot encoded as the task was classification. The color channels were not altered.

3.2. Deep learning modelling

This study employed Convolutional Neural Network (CNN)-based deep learning approaches for model building. Previous studies have shown that neural networks, such as CNNs, excel in analyzing image-based data by identifying both low-level and high-level features. Furthermore, we opted to use a technique called transfer learning. Transfer learning involves using an already pre-trained model to perform a new, similar task without having to start training the model from scratch. In this study, we trained and tested 4 popular transfer learning models -VGG16, ResNet-50, EfficientNetb7 and MobileNet. Figure 2 shows the VGG16 model architecture only as it was the best performing architecture in this work. VGG16 [21], a CNN model that was initially trained on the IMAGENET dataset, which contains over 14 million images classified into 22,000 categories.

Google collab was used to train and finally test all 4 models (names mentioned above) on 100 images. It provides free cloud computing as well as free processing power via graphics processing units (GPUs). Data data were reduced to a 0 to 1 scale. Gradient descent during during backpropagation, was used for optimisation of the models. The preprocessed image dataset was subjected to image augmentation techniques. Image augmentation allows to induce more variaiton in the dataset to meet unseen variaiton in the data sets. Augmentation makes the data more universal and helps the model comprehend it better. All four models were loaded in the google collab notebook with only last layers trainable. Activation function was set softmax and loss function to categorical cross entropy. Adam optimizer was selected with 0.01 learning rate. The batch size was set to 32 which means that during each run, 32 images were input to the model at a time for training.

Results and discussion

During training VGG16 achieved maximum stem node detection accuracy of 97%, as compared to other models trained. A summary of accuracy v/s epochs of different models trained is shown in Fig. 3. As can be noted, that for equal number of epochs, the learning behavior of VGG16 was the most stable as for other models the learning curve showed discontinuities as a function of epochs. Training for efficientnetb7 was the worst but this could be since it is a very lighter model compared to other models. The VGG16 also achieved better accuracy in the training with much less epochs compared to other models, indicating that VGG16 training was much more efficient than other models.

After model training, all four models were tested on 100 independent images, that were kept aside for testing purpose. VGG16 performed the best with 99% accuracy, followed by MobileNet – 90%, ResNet-50–83% and EfficientNetb7–52%. Table 1 shows confusion matrices of all 4 models after testing them on 100 images. As can be notes that for VGG16 only one samples was misclassified from 100 images, while for other there were major misclassifications.

Table 1
Summary of model testing performances (confusion matrices) on the independent test set for detecting stem node in sugarcane.

VGG16		
	Actual Positive	Actual Negative
Predicted Positive	49	0
Predicted Negative	1	50
MobileNet		
	Actual Positive	Actual Negative
Predicted Positive	42	2
Predicted Negative	8	48
ResNet-50		
	Actual Positive	Actual Negative
Predicted Positive	46	13
Predicted Negative	4	37
EfficientNetb7		
	Actual Positive	Actual Negative
Predicted Positive	52	48
Predicted Negative	0	0

Result of tests on 100 images (Table 1) directed us to select VGG16 for real time test and model deployment. A python graphical user interface program was developed and used to run real-time tests on the model. A setup comprising of a conveyor system (Fig. 4) and a stand for holding the mobile camera vertically above the conveyor belt was made. A mobile camera was used as a live streaming device for real-time testing. Five varieties of sugarcane (Co Pant 13224, Co Pant 97222, Co Pant 90223, Co Pant 12226, Co Pant 03220) having varying characteristics like color, girth diameter, internode length etc., were selected. Every frame of the live stream video was fed into the model, so it could only predict the node when it arrived at the area we wanted to test. The ROI [region of interest] concept was used in this case. The height of the cell phone camera was set to 22 cm and the forward speed of conveyor over which sugarcane was kept was set to 14 cm/sec. A 7-watt led bulb was used to lighting up the setup. The input video feed resolution was 640 x 480. Testing was done in the center of the video frame. A rectangular box was drawn at the middle region of video frame just to know where exactly the model was predicting, as shown in Fig. 5.

As mentioned earlier, the video resolution was set at 640 x 480 pixels. Instead of inputting the entire frame size into the model for predictions, which would have resulted in increased computational costs, prediction time, and slowed down the model's performance, we made the decision to extract a smaller frame size of 140 x 70 pixels from the center of the complete video frame. This approach effectively reduced computational expenses and allowed the model to make predictions more efficiently. The dimensions of the rectangle bounding box was the same as of the small frame size (140 x 70) which was fed into the model. Sugarcane nodes are detected as they move inside the box as shown in Fig. 6. The green circle on the left side of the image indicates the detection within the rectangular box.

Results of the real-time test conducted on five different sugarcane varieties are presented in Table 2. The table showcases the number of detected buds out of the total number of buds present on the sugarcane for each respective variety.

Table 2

Results of real time deployment and testing of VGG16 model to detect stem node. Note model was tested on different varieties and multiple runs were performed to check the stability of the model.

Variety – Co Pant 13224	
Total nodes present-9	
No. of runs	Nodes detected
1	8
2	8
3	9
Variety – Co Pant 97222	
Total nodes present-8	
No. of runs	Nodes detected
1	8
2	8
3	8
Variety – Co Pant 90223	
Total nodes present-8	
No. of runs	Nodes detected
1	8
2	8
3	8
Variety – Co Pant 1226	
Total nodes present-10	
No. of runs	Nodes detected
1	10
2	10
3	9
Variety – Co Pant 03220	

Variety – Co Pant 13224	
Total nodes present-13	
No. of runs	Nodes detected
1	11
2	12
3	10

The average detection time for each frame was 0.182 seconds. One of the key advantages of this computer vision system is its high generalizability to a wide variety of sugarcane strains as shown while testing the model on five different varieties of sugarcanes. This is because CNNbased models can capture numerous independently identifiable features of an object, such as its color, shape, size, and pattern. All that is required is high-quality image data that has been captured in the appropriate settings for deploying the model. While the Yolov3 models were also viable options, they are more computationally expensive than VGG16, making it less expensive to operate on single-board computers. By implementing this approach, automatic sugarcane cutting machines can produces results that are on par with, or even better than, existing object detection/localization algorithms. In our model, we can specify the frame area where we want the model to search for the presence of a node.

Conclusion

The findings of the study indicate that computer vision and convolutional neural networks (CNN) were successful in detecting sugarcane stem nodes. The validation of the CNN model on different sugarcane varieties demonstrated a high accuracy rate of 99% in recognizing sugarcane stem nodes. Among the four compared CNN models, VGG16 exhibited superior performance by learning more information in fewer epochs and achieving higher accuracy in stem node detection during both offline and inline deployment tests. Furthermore, the detection process could be conducted in real-time video, taking only 0.182 seconds, which makes it suitable for inline applications of computer vision in stem node detection.

This cost-effective approach has the potential to be integrated into sugarcane stem node cutting/chipping machinery, with some modifications and appropriate hardware choices, to achieve higher detection performance. The proposed methodology can also be employed to precisely cut sugarcane sets without causing any bud damage in sugarcane planters. By doing so, it can significantly reduce labor requirements and save time, providing substantial benefits to end-users. Moreover, the deep transfer learning methodology presented in this study can be adapted to other applications within the agricultural domain. The primary advantage of deep transfer learning is the ability to train successful models with minimal sample points for advanced computer vision tasks, such as stem node detection, as demonstrated in this study.

Declarations

Competing Interests : The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Information : No funding received.

Author contribution :

Arun Kumar: Conceptualization, Methodology, Investigation, Writing - Original Draft, Project administration

Abhishek Kumar Saini: Methodology, Software, Investigation, Writing - Original Draft

Alaknanda Ashok: Methodology, Investigation, Writing - Original Draft

Puneet Mishra: Writing - Original Draft, Writing - Review & Editing

References

1. Solomon, S.: Sugarcane Production and Development of Sugar Industry in India. Sugar Tech. **18**(6), 588–602 (2016)
2. Dinesh Babu, K.S., et al.: A short review on sugarcane: its domestication, molecular manipulations and future perspectives. Genet. Resour. Crop Evol. **69**(8), 2623–2643 (2022)
3. Gill, J.S., Kaur, G.: Infusion of single bud chip planting technique for sugarcane propagation. Indian J. Econ. Dev. **11**(1), 227–232 (2015)
4. De Ketelaere, B., et al.: A fresh look at computer vision for industrial quality control. Qual. Eng. **34**(1), 152–158 (2022)
5. Tian, H., et al.: Computer vision technology in agricultural automation —A review. Inform. Process. Agric. **7**(1), 1–19 (2020)
6. LeCun, Y., Bengio, Y., Hinton, G.: Deep Learn. Nat. **521**(7553), 436–444 (2015)
7. Marini, F.: Artificial neural networks in foodstuff analyses: Trends and perspectives A review. Anal. Chim. Acta. **635**(2), 121–131 (2009)
8. Alzubaidi, L., et al.: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J. Big Data. **8**(1), 53 (2021)
9. Polder, G., et al.: *Potato Virus Y Detection in Seed Potatoes Using Deep Learning on Hyperspectral Images*. Front. Plant Sci., 10. (2019)
10. Polder, G., et al.: *Automatic Detection of Tulip Breaking Virus (TBV) Using a Deep Convolutional Neural Network This work was partially funded by the Dutch Product Board for Horticulture and the Dutch Ministry of Economic Affairs, Agriculture and Innovation. The research is part of the public*

private partnership BollenRevolutie 4.0 funded under the grant TKI-TU-1806. IFAC-PapersOnLine, 52(30): p. 12–17. (2019)

11. Mahlein, A.K., et al.: Hyperspectral Sensors and Imaging Technologies in Phytopathology: State of the Art. Annu. Rev. Phytopathol. **56**(1), 535–558 (2018)
12. Lu, S., et al.: Recognition and features extraction of sugarcane nodes based on machine vision. Nongye Jixie Xuebao = Transactions of the Chinese Society for Agricultural Machinery. **41**(10), 190–194 (2010)
13. Wei-zheng, Z., et al.: *Research on identification and location method of sugarcane node based on hyperspectral imaging technology*. J. Light Ind., **32**(5). (2017)
14. Zhou, D., et al.: *Identification and Localisation Algorithm for Sugarcane Stem Nodes by Combining YOLOv3 and Traditional Methods of Computer Vision*. Sensors, **22**(21). (2022)
15. Zhou, D., et al.: A new design of sugarcane seed cutting systems based on machine vision. Comput. Electron. Agric. **175**, 105611 (2020)
16. Chavan, P., et al.: AUTOMATIC SUGARCANE NODE CUTTING MACHINE. International Journal of Innovative Research in Science and Technology (2016)
17. Rattanaphongphak, A., Boongsood, W.: Design of machine vision system for sugarcane buds or rings detection. J. Image Graphics. **7**(3), 102–106 (2019)
18. Chen, M., et al.: Sugarcane Stem Node Detection Based on Wavelet Analysis. IEEE Access. **9**, 147933–147946 (2021)
19. Xu, J., Mishra, P.: *Combining deep learning with chemometrics when it is really needed: A case of real time object detection and spectral model application for spectral image processing*. Anal. Chim. Acta, : p. 339668. (2022)
20. Mishra, P., Passos, D.: Realizing transfer learning for updating deep learning models of spectral data to be used in a new scenario, p. 104283. Chemometrics and Intelligent Laboratory Systems (2021)
21. Simonyan, K., Zisserman, A.: *Very deep convolutional networks for large-scale image recognition.* arXiv preprint arXiv:1409.1556, (2014)

Figures

Figure 1

A sample node (left) and internode (right) image captured by the camera.

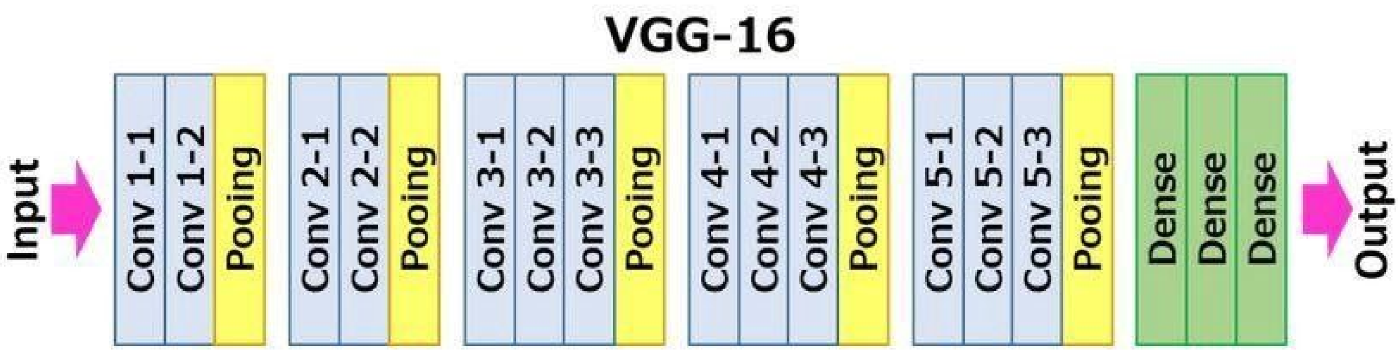

Figure 2

A summary of different model layer for VGG16 deep learning architecture.

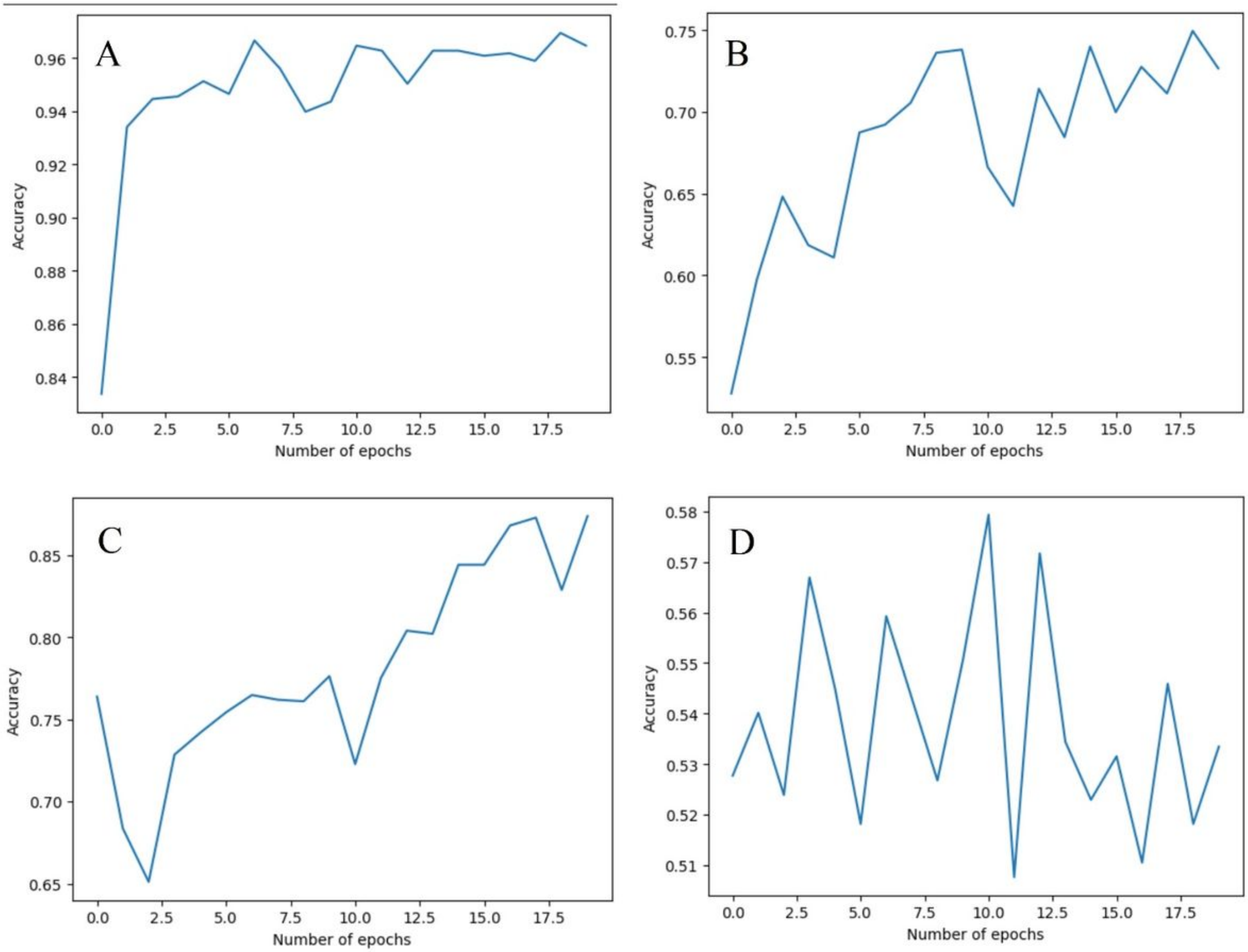

Figure 3

Accuracy v/s epochs graphs. (A- VGG16, B- ResNet50, C- MobileNet, D- EfficientNetb7)

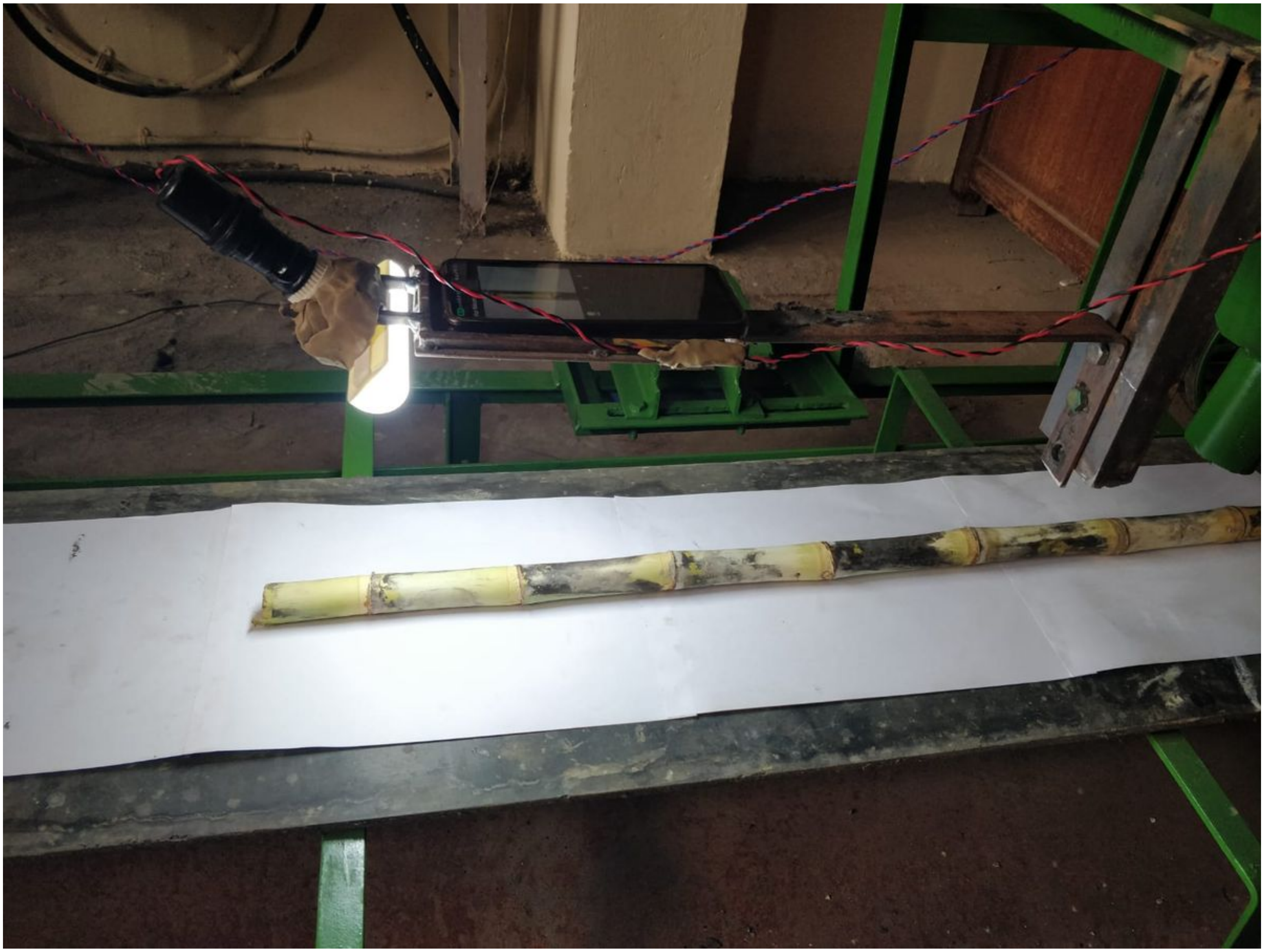

Figure 4

Model system for real-time model deployment for stem node detection based on VGG16 trained model.

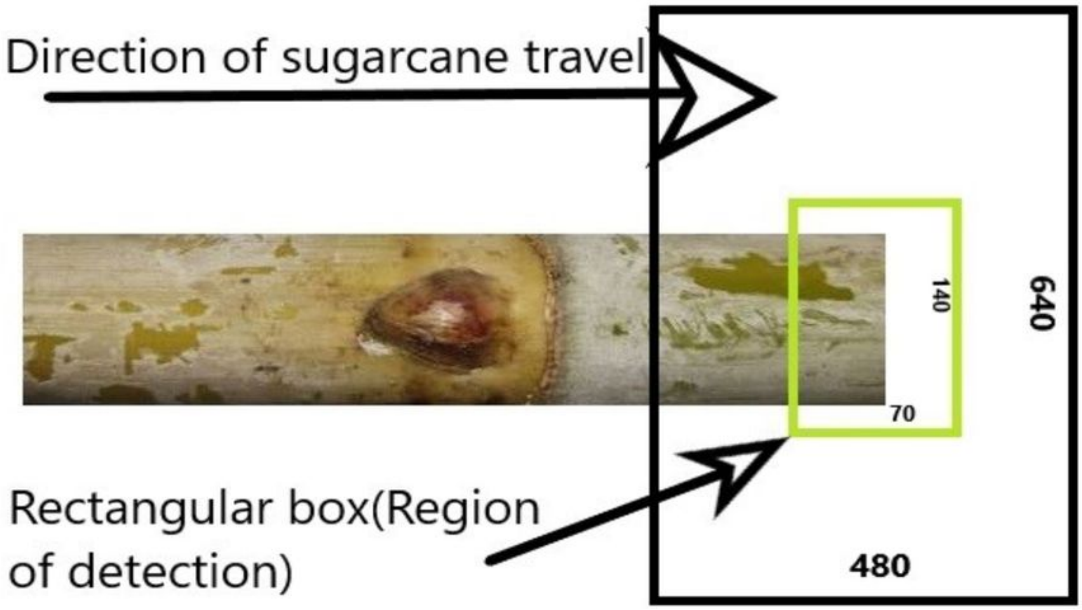

Figure 5

The schema for rectangular bounding box detection in live stream video.

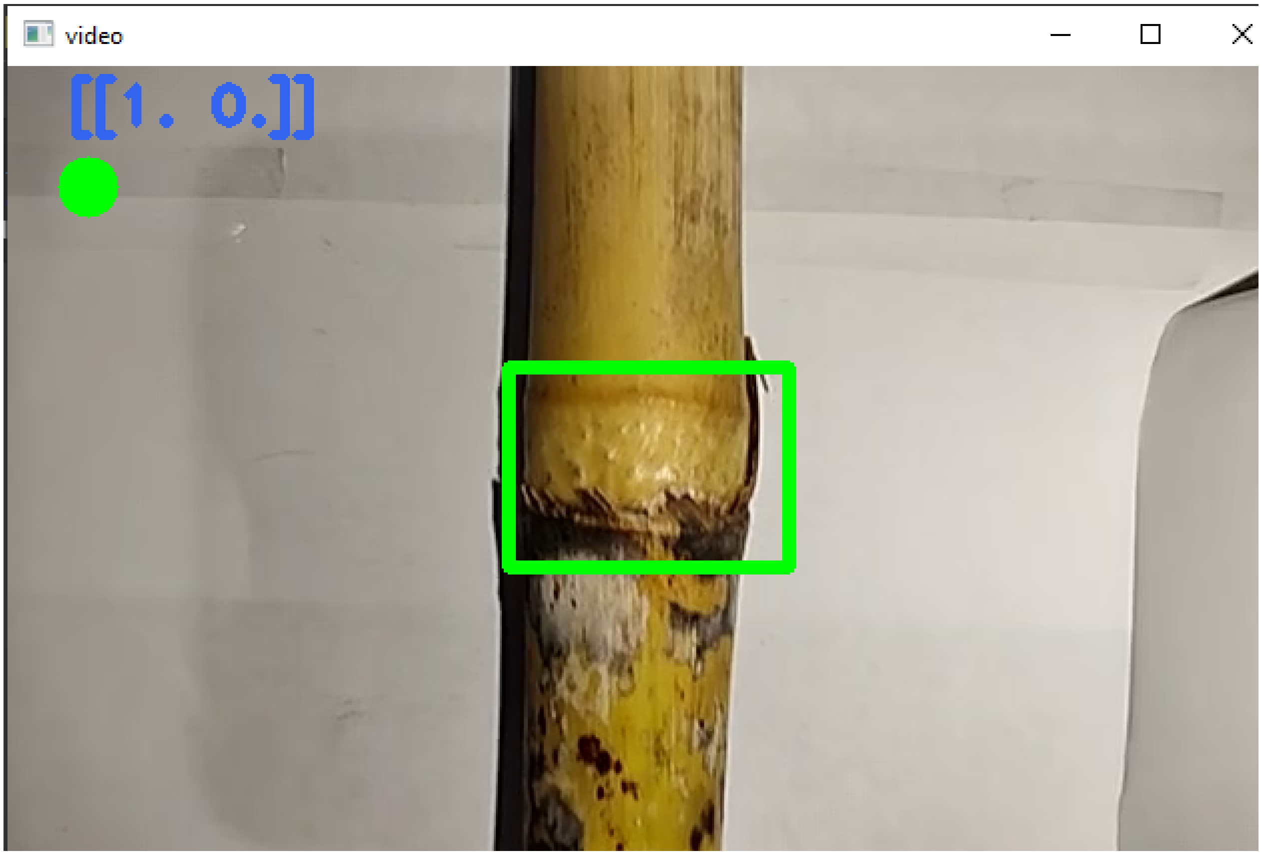

Figure 6

Detection of stem node in real time with the VGG16 model and the python-based deployment application developed in this study.