International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Grape Vision: A CNN-Based System for Yield Component Analysis of Grape Clusters

[1] Dr. B. J. Dange, [2] Dr. Punit Kumar Mishra, [3] Kalpana V. Metre, [4] Santosh Gore, [5] Dr. Sanjay Laxnamrao Kurkute, [6] Dr. H. E. Khodke, [7] Sujata Gore

Submitted: 27/04/2023 **Revised**: 24/06/2023 **Accepted**: 05/07/2023

Abstract*:* The agricultural industry is adopting advanced technologies and applications like yield prediction, precision agriculture, and automated harvesting to enhance production and quality. Machine learning (ML) and computer vision are increasingly used for fruit detection, segmentation, and counting. Specifically, the use of Convolutional Neural Networks (CNN) in grape yield prediction and quality assessment is gaining popularity due to its high accuracy and cost efficiency. Additionally, a new methodology based on image analysis has been developed for fast and inexpensive cluster yield component determination in the wine and table grape industry.

Keywords: *Grape prediction, CNN, machine learning, image processing*

1. Introduction

The agricultural industry has recently undergone a significant upgrade through the implementation of automation, particularly in the field of grape cultivation. To achieve optimal results, it is crucial to keep up-to-date with the latest advancements in research related to this field. Artificial intelligence (AI)-based techniques are now being utilized to support predictions, thereby replacing manual estimation methods [1].

Grapes are a vital crop that supports the economic progress of farmers. Therefore, the automation of grape fields plays a crucial role in the enhancement of monitoring and quality prediction of grapes to attain quality standards. The quantity of grapes produced is not as important as their quality. Hence, the automation of grape monitoring can improve output while increasing the quality of grapes. This approach also benefits farmers by reducing production costs and minimizing the environmental impact of grape cultivation [1].

Grapes are particularly significant in terms of improving output when compared to other fruits. However, because it depends on several variables, forecasting fruit quality can be difficult. These elements include outward characteristics of the fruits, such as their size, colour, and other characteristics. Therefore, it is crucial to use cutting-edge technology like AI to forecast crop quality to increase the quality of grapes. Farmers can increase their income while ensuring that their grape output adheres to good standards by doing.

Fruit attributes like size and colour are frequently analysed using computer vision techniques. Apples, bananas, and grapes are just a few examples of fruits whose colours are frequently thought to signify a degree of ripeness. However, in addition to these visual characteristics, such as size, texture, and form, grape ripeness is also determined by additional factors [2].

Two main ways to analyse grape qualities are distinguished using machine learning techniques [3]. In the first method, input grape photos are applied to a model that has been fully trained. The second method, referred to as transfer learning, is analysing similar problems by applying information from problems that have already been resolved to new ones. This method involves using a pre-trained network to extract features from a related task and using that information to improve a new classifier [4].

2. Related Work

Recent research has revealed that ML models can more accurately categorize grape bunches during the medium phase of development, compared to the phase after bloom.

[1]*Computer Engineering department, Sanjivani College of Engineering Kopargaon, (An Autonomous institute), Affiliated to Savitribai Phule Pune University Pune, Maharashtra, India, bapudange@gmail.com.*
[2]*Marketing Management, Symbiosis Centre for Management Studies, Symbiosis International (Deemed University) Pune, Maharashtra, India, punit.mishra@scmspune.ac.in.*
[3]*MET's Institute of Engineering, Nashik, Maharashtra, India, kvmetre@gmail.com.*
[4]*Director, Sai Info Solution, Nashik, Maharashtra, India, https://orcid.org/0000-0003-1814-591, sai.info2009@gmail.com.*
[5]*Savitribai Phule Pune University, Pravara Rural Engineering College Loni, Maharashtra, India ,drsanjaykurkute@gmail.com.*
[6]*Assistant Professor, Computer Engineering department, Sanjivani College of Engineering Kopargaon, (An Autonomous institute), Maharashtra,India,423603, Savitribai Phule Pune University, Pune, India, hekhodke@gmail.com.*
[7]*Director, Sai Info Solution, Nashik, Maharashtra, India sujatarpatil21@gmail.com.*

This is attributed to the smaller size of grape bunches in the latter phase. However, identifying grape bunches that are identical in colour and texture to their surrounding flora can be quite challenging. Therefore, implementing advanced viticulture process management techniques, such as computer vision-based crop monitoring systems, is essential for the automatic supervision of crops. In addition to using computer vision, remotely identifying information with the help of unmanned vehicles is also crucial. As emphasized in [4], improving the control of unmanned vehicles is a critical task for the enhancement of remotely identified information. By doing so, viticulturists can better manage their crops and ensure optimal grape production. Furthermore, using advanced monitoring systems can help to identify any issues early on, allowing for timely intervention to prevent significant crop loss.

Accurately predicting vineyard production levels in different regions is critical to achieving optimal harvesting techniques in viticulture. According to a study on viticulture [5], computer vision techniques have vast potential in the precise categorization of crops. These techniques can differentiate between the canopy and soil in images, thereby reducing failure rates in grape detection and providing better results with higher accuracy. As a result, the implementation of computer vision techniques can significantly improve the overall quality of grape production in viticulture.

In recent studies, the use of CNNs has proven to be both effective and efficient in implementing robust grape prediction systems. However, the process of feature engineering has shifted towards acquiring more effective representations [6] in ML systems. This approach has been successful in overcoming issues associated with grape prediction. In addition to CNNs, deep learning has been introduced to broaden research in the agricultural field. To solve agricultural challenges, this method uses an image-based perceptual approach. Deep learning is becoming more and more common in agricultural applications, like yield estimation, according to the author of [7]. In the following years, deep learning is anticipated to improve to the point where it is a necessary instrument for the development of agriculture.

CNNs are frequently employed in visual feature-based computer vision systems to extract features from the input visual data [8]. However, the classification of berries can be influenced by their location [9], which may vary. Nevertheless, with sufficient training data samples, CNNs can effectively predict these differences, including variations in the posture, colour, and light intensity of the berries. Therefore, it is crucial to evaluate the quality and variability of training data when assessing input imaging data for pixel classification in CNN and fruit segmentation [10].

In recent research, end-to-end grape detection using faster R-CNN techniques has become more popular than conventional techniques, as it has been found to improve grape detection efficiency [11] [12]. These advanced techniques offer robust and efficient systems for grape detection and have shown promising results in the field of agricultural automation.

Nuske et al. [13] have demonstrated the feasibility of accurate field processes using computer vision techniques without requiring background control. Their proposed approach employs the Radial Symmetry Transform method proposed by [14] for multistage berry candidate detection. The K-nearest neighbours (KNN) classifier is then used to filter features such as the colour and texture of the berries, and isolated berries are removed to form groups of berries. Their model achieved a precision rate of 98% for three grape varieties using a dataset of 2,973 berries and a recall rate of 63.7. Moreover, their approach showed a correlation score of 0.74 between the berry count for each vine and its corresponding weight. These findings highlight the potential of computer vision techniques to improve the accuracy and efficiency of grape detection in the agricultural industry.

Fig. 1. Field process of grape yield

3. Methodology

Efficient object classification using ML techniques has gained significant attention in recent years. In a study aimed at improving grape detection and classification, a CNN-based model was proposed. The model was trained with a real-world image dataset, which included various grape bunches with different sizes, shapes, and colours.

The success of ML models heavily relies on the quality and quantity of data used for training. Therefore, it is crucial to select an efficient dataset and properly preprocess it before training the model. In the case of grape detection and classification, a comprehensive dataset can include different grape varieties, growth stages, and lighting conditions to ensure the model's robustness and generalization capability.

The proposed CNN-based model aims to overcome the limitations of traditional methods, such as manual labour and low efficiency, in grape detection and classification. By leveraging the power of deep learning and computer vision, the model can accurately identify and classify grape bunches in real-time, facilitating automatic crop supervision and improved crop management.

1.1. Image processing

The GrapeVision system is a CNN-based method that makes use of image processing to examine grape cluster yield components. Image acquisition, image processing, and yield component analysis are the system's three main phases. A high-resolution camera is used to take pictures of grape clusters during the image acquisition phase. After that, any noise or distortion in the photos is removed during preprocessing. The preprocessed photos are segmented during the image processing stage to remove the backdrop and reveal the grape clusters. Edge detection methods, morphological processes, and colour thresholding are combined to achieve this. Following the segmentation of the grape clusters, CNN-based feature extraction and classification approaches are used to analyse several yield components, including berry count, berry size, and cluster weight. The system achieves high accuracy in yield component analysis, demonstrating the potential for image processing to play an important role in precision agriculture for grape production.

1.2. CNN

Artificial neural networks called CNNs are widely used in computer vision due to their high performance [15]. They can efficiently perform complex classification tasks by learning and extracting features from end to end [16]. However, to capture the diverse features from images, CNNs require a large number of training data samples. These training images must encompass all the variations of grapes that may be encountered during the test phase, to provide features of different geometric transformations.

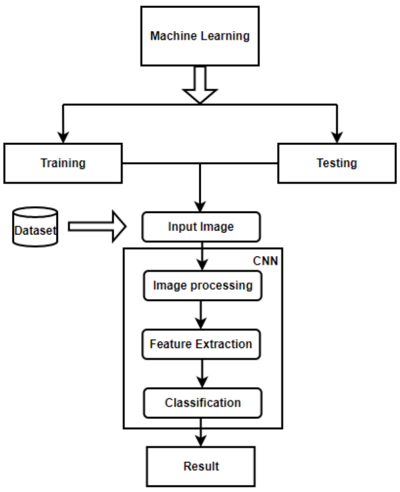

Fig. 2. Working flow of the proposed model

CNNs can identify individual grape bunches and recognize fruit boundaries through image processing techniques such as filtering and segmentation [17]. While circular grape shapes can be easily estimated for size and shape, adjusting to rectangular boxes can be challenging. Instance segmentation, as shown in Figure 2, is a precise technique that allows for pixel classification and identification of non-grape components such as leaves, branches, and trunks [17]. The feature extraction process in CNN can classify images as grape and non-grape components, and the grape components can be analyzed for defined features to track the growth of bunches during the initial, intermediate, and final stages.

To ensure the precise training of the model, it is tested with a variety of grape image datasets and a validation dataset is employed with selected validation images representing the complete dataset and its features. The proposed model is validated with a significant number of images that include a variety of grapes, and it fits well with a fully trained network. The performance of the model is verified by comparing real-world images with trained images.

1.3. Dataset

To identify grape bunches of varying sizes, shapes, colours, and textures, a set of training images is necessary. These feature parameters are crucial for differentiating between various grape varieties and grapes of the same variety. The training, validation, and test sets are created by randomly assigning images in an 80-20% proportion to each set. The training dataset images are masked to isolate the background and obstruct foreground pixels.

4. Result

The proposed ML model demonstrates efficient performance in predicting grape development across a variety of grape images. The size and colour of grapes in each developmental phase are critical factors for accurate prediction. Performance evaluation metrics, such as accuracy, precision, and recall, are used to assess the effectiveness of grape prediction in the proposed model.

In addition, various data mining algorithms have been explored to compare their effectiveness in grape classification using images. Support Vector Machines (SVM) and Multinomial Logistic Regression (MLR) are studied and compared with Random Forest (RF) and K-Nearest Neighbors (k-NN) algorithms.

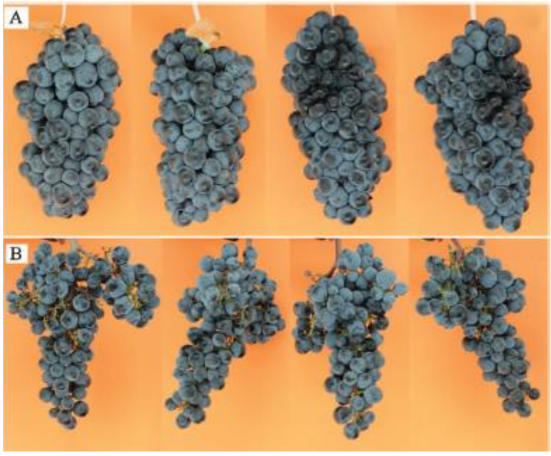

Fig. 3. Four views of grape bunch detection

TABLE I. PERFORMANCE OF MODEL FOR TRAINING AND TESTING

Algorithm	Sensitivity	Testing Accuracy	Specificity	Training Accuracy
Support Vector Machine	95	87.9	86	85.8
Multinomial Logistic Regression	86	79	89	88.7
K-Nearest Neighbour	90	84.7	99.99	84.6
Random Forest	83	94.5	97	95.6

The ML model's performance was evaluated using various algorithms such as RF, MLR, SVM, and KNN. The results from Table I indicate that the RF algorithm achieved the highest training and testing accuracy. However, the MLR algorithm was more efficient in training but less efficient in testing. The SVM and KNN algorithms had moderate performance in predicting grape quality at different developmental stages. To determine the most relevant technique, the specificity and sensitivity of each algorithm were compared. KNN had the highest specificity, followed by RF, MLR, and SVM. In contrast, SVM had excellent sensitivity. Figure 4 provides a graphical representation of the training and testing accuracy for each algorithm. These results demonstrate the importance of selecting the appropriate algorithm based on the specific application and performance criteria.

TABLE II. PERFORMANCE OF GRAPE PREDICTION

Variety of grapes	Parameter		
	Accuracy	***Precision***	***Recall***
Variety 1	90.8	92	94.5
Variety 2	91.8	96	97
Variety 3	90	97	98
Variety 4	90	95	96
Variety 5	91.4	92	94

The accuracy of the ML model for five different grape varietals at various phases of growth is presented in detail in Table II. The accuracy range of the grape quality prediction model is between 90% and 92%, demonstrating a good level of performance. The table gives a fuller picture of the model's predictive skills by including precision and recall values for each grape variety in addition to accuracy numbers. Figure 5 presents each grape variety's performance parameters for a more thorough knowledge of the model's performance.

5. Discussion

The suggested approach uses image processing methods to recognise grape clusters and monitor their growth in the early, middle, and late stages. For visual recognition of grape clusters, the model contains several components. Additionally, during training and testing, the performance of the algorithms RF, MLR, SVM, and KNN are used to assess the performance of the grape quality prediction. On five different grape kinds, each of which includes the starting, intermediate, and final stages of harvesting, a pre-trained model is put to the test. The model makes precise predictions for each stage depending on the growth of the grapes.

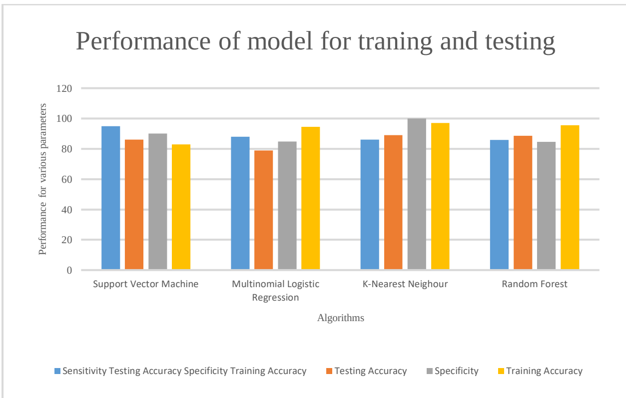

Fig. 4. Training and testing accuracy of the grape prediction model

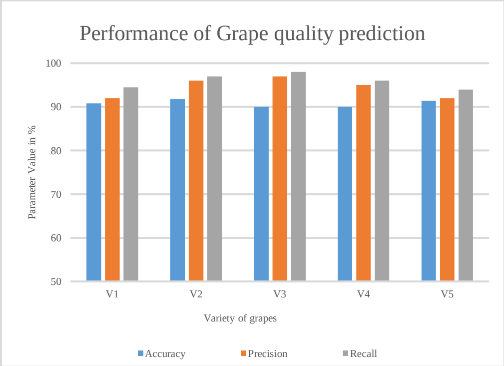

Fig. 5. Performance values gained from the Grape quality prediction model

The focus of our study is on the size of grapes as it increases during development. By evaluating the size of grapes in images, we predict the quality of grapes at different stages of harvesting. The role of the classifier in achieving the classification of grape clusters is critical. Our CNN model results demonstrate significant identification of grapes in each development stage, as well as the detection of grape varieties that differ in size and shape. The ML model classifies grapes into three different stages based on their size, with larger sizes indicating higher quality grapes at the harvesting stage. Our model has achieved an accuracy above 90%. The efficiency of the model depends on the features extracted from the images, and having a dataset with multiple varieties of features can increase its performance.

6. Conclusion

Our research presents a novel approach to grape quality prediction using a CNN-based model that accurately assesses the quality of grapes at various growth stages and sizes, taking into account colour as an additional indicator. Our model extracts features from grape images using CNN, considering the variety of grape features that may vary across different grape types. Furthermore, we evaluate the performance of our model on five grape varieties and compare different classifiers to achieve the highest accuracy. The proposed method provides a promising solution for the development of intelligent agri-tech systems for smart vineyards, enabling accurate and efficient prediction of grape quality based on growth and size.

Reference

[1] A. S. Aguiar et al., “Grape bunch detection at different growth stages using deep learning quantized models,” Agronomy, vol. 11, no. 9, 2021, doi: 10.3390/agronomy11091890.

[2] K. Kangune, V. Kulkarni, and P. Kosamkar, “Grapes Ripeness Estimation using Convolutional Neural network and Support Vector Machine,” in 2019 Global Conference for Advancement in Technology, GCAT 2019, 2019. doi: 10.1109/GCAT47503.2019.8978341.

[3] J. L. Aleixandre-Tudo, A. Buica, H. Nieuwoudt, J. L. Aleixandre, and W. Du Toit, “Spectrophotometric Analysis of Phenolic Compounds in Grapes and Wines,” J. Agric. Food Chem., vol. 65, no. 20, pp. 4009–4026, May 2017, doi: 10.1021/ACS.JAFC.7B01724.

[4] L. Comba, A. Biglia, D. Ricauda Aimonino, and P. Gay, “Unsupervised detection of vineyards by 3D point-cloud UAV photogrammetry for precision agriculture,” Comput. Electron. Agric., vol. 155, pp. 84–95, 2018, doi: 10.1016/j.compag.2018.10.005.

[5] M. Sozzi, S. Cantalamessa, A. Cogato, A. Kayad, and F. Marinello, “22. Grape yield spatial variability assessment using YOLOv4 object detection algorithm,” 2021, pp. 193–198. doi: 10.3920/978-90-8686-916-9_22.

[6] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 8, pp. 1798–1828, 2013, doi: 10.1109/TPAMI.2013.50.

[7] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” Computers and Electronics in Agriculture, vol. 147. pp. 70–90, 2018. doi: 10.1016/j.compag.2018.02.016.

[8] I. Goodfellow, Y. Bengio, and A. Courville, “Deep Learning,” MIT Press (online available), 2016, Accessed: Jan. 12, 2023. [Online]. Available: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Goodfellow%2C+I.%2C+Bengio%2C+Y.%2C+Courville%2C+A.%2C+2016.+Deep+Learnin&btnG=

[9] S. Nuske, K. Wilshusen, S. Achar, L. Yoder, S. Narasimhan, and S. Singh, “Automated visual yield estimation in vineyards,” J. F. Robot. , vol. 31, no. 5, pp. 837–860, 2014, doi: 10.1002/rob.21541.

[10] S. W. Chen et al., “Counting Apples and Oranges with Deep Learning: A Data-Driven Approach,” IEEE Robot. Autom. Lett., vol. 2, no. 2, pp. 781–788, 2017, doi: 10.1109/LRA.2017.2651944.

[11] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.

[12] X. Liu et al., “Monocular Camera Based Fruit Counting and Mapping with Semantic Data Association,” IEEE Robot. Autom. Lett., vol. 4, no. 3, pp. 2296–2303, 2019, doi: 10.1109/LRA.2019.2901987.

[13] S. Nuske, S. Achar, T. Bates, S. Narasimhan, and S. Singh, "Yield estimation in vineyards by visual grape detection," in IEEE International Conference on Intelligent Robots and Systems, 2011, pp. 2352–2358. doi: 10.1109/IROS.2011.6048830.

[14] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 8, pp. 959–973, 2003, doi: 10.1109/TPAMI.2003.1217601.

[15] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 3642–3649, 2012, doi: 10.1109/CVPR.2012.6248110.

[16] L. Zheng, Y. Yang, and Q. Tian, "SIFT Meets CNN: A Decade Survey of Instance Retrieval," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 5. pp. 1224–1244, 2018. doi: 10.1109/TPAMI.2017.2709749.

[17] T. Y. Lin et al., "Microsoft COCO: Common objects in context," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2014, vol. 8693 LNCS, no. PART 5, pp. 740–755. doi: 10.1007/978-3-319-10602-1_48.

[18] Yathiraju, D. . (2022). Blockchain Based 5g Heterogeneous Networks Using Privacy Federated Learning with Internet of Things. Research Journal of Computer Systems and Engineering, 3(1), 21–28. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/37

[19] Kalyani, B. ., Sai, K. P. ., Deepika, N. M. ., Shahanaz, S. ., & Lohitha, G. . (2023). Smart Multi-Model Emotion Recognition System with Deep learning. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1), 139–144. https://doi.org/10.17762/ijritcc.v11i1.6061