Mango not only contains fundamental dietary nutrients but is an essential source of income in the Gambia. Unfortunately, farmers are unable to utilize its complete production due to among other things, the lack of technology that would allow them estimate production in advance. In an attempt to fill this gap, the YOLOv5 framework is used to detect a variety of mangoes in the Gambia. We implemented this by modifying the YOLOv5 structure to build a new backbone using Conv2D + Batch Normal + Swish (CBS). We further performed experiments by comparing our proposed system with models of existing frameworks such as Faster RCNN,YOLOv3,YOLOv4, EfficientDet. Experimental results show that our model performs relatively better compared to the other four models.

Introduction

• Mango is one of the main fruit produced, consumed, and exported in The Gambia. Currently, mango farmers incur many losses because there are no non-destructive ways for determining fruit output.

• Knowledge of the production of mango fruits can be a vital means to implement trading strategies. Taking a proper count of the number of mango fruits in an orchard leveraging deep learning techniques would be one of the most reliable methods of acquiring production statistics [1].

• In [2], the new version of Faster R-CNN and YOLOv3 are employed for the detection and comparison of oranges, mangoes, and apples on the same dataset. Experimental results reveal that YOLOv3 is at most 4% better than the usual Faster R-CNN through the recognition of roughly 90% of fruits. YOLOv3 had a 40ms average recognition time, compared to 58ms of the improved Faster R-CNN algorithm and 240ms for conventional Faster RCNN.

Methodology

Our System works by first getting video frames of mangoes from a camera feed. This frames are then fed into our Proposed YOLOv5 network and then inference is performed to produce an output of mangoes detected.

• For our experimental results, we use the mAP and FPS as the metrics to measure our models’ accuracy and inference time.

• Fig. 2 shows our YOLOv5 network with Conv2D + Batch Normal + Swish (CBS) module.

• We evaluated the proposed methodology’s performance using the accuracy and loss of our presented model and obtained a validation accuracy of 94.51 and a validation loss of 0.0273.

• We used the transfer learning technique and trained a different model on our generated dataset to compare with other state-of-the-art frameworks. As shown, our proposed model outperformed the rest of the models.

Conclusion

We introduce in detail the YOLOv5 method of detecting objects, in our case to count mango fruits per tree. We also discuss the architecture of the YOLOv5 network, the setup of our proposed classifier, and dataset processing. Results show that the YOLOv5-based method of detecting mangoes achieved outstanding results in terms of speed and accuracy. We also generated a dataset for a local environment that is publicly available for researchers.

References
