Qualiphal - Computer Vision based Quality Assurance system for supply chain of fruit delivery and import/export.

Abhimanyu Banerjee
Computer Science and Engineering
Bennett University
Gurgaon, India
Ab8963@benett.edu.in

Anudit Nagar
Computer Science and Engineering
Bennett University
Noida, India
An9316@bennett.edu.in

Himanshu Mittal
Computer Science and Engineering
Bennett University
Bahadurgarh, India
Hm6729@bennett.edu.in

***Abstract*—This paper attempts to provide an effective solution for quality assessment of fruits that undergo deterioration during the process of transportation and delivery. Presently used manual methods of quality checking are labour intensive. A viable computer vision based alternative will help in automation of this process. In contrast to classical machine learning based approaches we propose to use high performing deep learning architectures like the U-Net and FPN (Feature Pyramid Network) for the purpose of multi class segmentation. We found that dilated FPN performs better segmentation on tiny regions compared to UNet in the lemon dataset which is common in fruit quality segmentation tasks and thereby producing quite satisfactory results. This will help ensure fresh fruits are reaching the consumers from the farms, wholesale markets/mandis and eliminate fruits that undergo deterioration during any stage of the supply chain.**

***Index Terms*—Computer Vision, Supply Chain Optimization, Image Processing, Deep Neural Networks.**

I. Introduction

Fruit exports in India amounted to nearly 764 million US dollars in the fiscal year 2020. This industry seems to be expanding over the years.So there is a need to keep up with the quality standards of the fruits being exported out of the country.Quality Testing of fruits and vegetables is an important aspect in maintaining the health of the entire nation. With increase in consumers across the world that are now engaging in more and more online doorstep delivery of fruits and vegetables especially during the hard times of COVID-19, the food industry needs to keep up with the quality of delivered fruits and vegetables.

The application of artificial intelligence in agriculture is increasing substantially in recent years as it provides substantial insight towards analysing patterns between their behaviours and how their quality gets impacted. Quality testing of edibles is done via examining various factors like color, shape, size, texture, and time since harvest. The spectrum of damages that these fruits and vegetables can endure during their life cycle is huge. Our purpose is to study techniques to understand these patterns and develop efficient methods to classify these fruits and vegetables in a non-invasive and non-destructive fashion. By studying the effects of these damages endured by the fruits and vegetables we present various feature-extraction methods for classification tasks.

In contrast to traditional techniques for fruit quality assessment such as spectroscopy and various sensing techniques, we are attempting to come up with a computer vision based cost-effective solution that will be a promising tool for quality assurance in the food industry. This can help reduce costs and improve longevity of the fruits and vegetables during long travels from farms to last mile delivery networks.

Also based on initial assessment of quality, we can determine which all fruits must be supplied first out of the wholesale storage and silos before quality becomes even lower and eventually leading in wastage.Almost 30-35 % of the vegetables and fruits perish at various stages. Our study will help reduce wastage of food by analyzing the cause and identification of ill regions on the surface of a fruit therefore reducing wastage and increasing the amount of value added products in the market.

Most works attempted till now in this domain of quality assessment have used classical approaches like edge detection, clustering techniques[13] and Bayesian classifier. Through this paper we propose deep learning methods for image segmentation using high yielding architectures like U-Net. Read section III for detailed information.

II. Related Work

Research has now paved its way into food sciences and engineering looking for advanced and efficient techniques for quality assurance . The traditional methodologies of dielectric spectroscopy and sensing[8, 14] which make use of sophisticated instrumentation and measurement techniques make the process of fruit quality evaluation difficult to implement in the supply chain mechanism. To cater to the industrial need of alternative measures, a machine learning based approach was experimented.

The analysis of fruit quality in the machine learning aspect makes use of image processing techniques [6]. The captured fruit images undergo pre-processing to remove noise from the images and prepare the data set for further inspection.

Computer vision and segmentation technology is further used to classify the fruit into defected and non defected categories [5]. Based on external features like shape, colour, blemishes and few other external features classification of the fruit image is carried out[4, 3, 22]. Detailed work [11] on feature extraction and learning methods like K Nearest neighbours, support vector machines and neural networks (ANN and CNN) have been covered in the referenced papers.

Aiming to reduce the manual work of sorting defected or low quality food, image classification and segmentation are used to comprehend the fruit images[18]. The paper published by Dr. Alam and co-authors proposed a quality assessment system[12] for rice grains using optical image processing using a number of characteristic features. Similar such work has been done in recognizing and classifying infected apples [21] as well as for identifying ripe coffee fruits[17].

With increasing demand for automation in the agricultural sector, fruit and vegetable grading and sorting systems have been developed using segmentation techniques such as Otsu method[7]. Semantic segmentation which is nothing but pixel level prediction and classification grouped with the high yielding convolutional networks gives a boost to image classification and recognition. [15, 9, 10] Prior work implementing convnets for semantic segmentations have been attempted in which every pixel is labelled with a class and our paper extends their work with further experimentations and improvisation.

Advanced architectures such as the U-Net are used to get accurate segmentation results in the medical field [16]. The work done in this direction for diagnosis purpose with the help of radiographic images, CT scans and MRIs can also be proliferated in food sciences. Such high performing convolutional networks which provide highly precise localization in biomedical image segmentation [20] can be extended in our proposed work to create a large scale automated fruit quality evaluation system and and bring about quality assurance in the supply chain.

III. Methodology

In this section we hereby elaborate on the specifics of the acquired dataset for training and testing as well as the details of the methodologies and approaches incorporated in this paper.

A. Data Acquisition

Acquiring quality data is of utmost importance for any model to learn and perform effectively. We came across various data sets like that of papaya and citrus fruits, but out of all of them we decided to go with the dataset of lemons gathered by 'SoftwareMill' [2]. The lemon dataset 1 being well labelled made it much better suited for our objective of multi class segmentation. The dataset comprises a total of 2690 annotated images, each with the dimensions of 1056 x 1056 pixels. The data has been annotated by broadly 7 labels namely- illness,gangrene, mould, blemish, dark style remains, pedicel, artifact. This comprehensive dataset is available to us in COCO format containing JSON files.

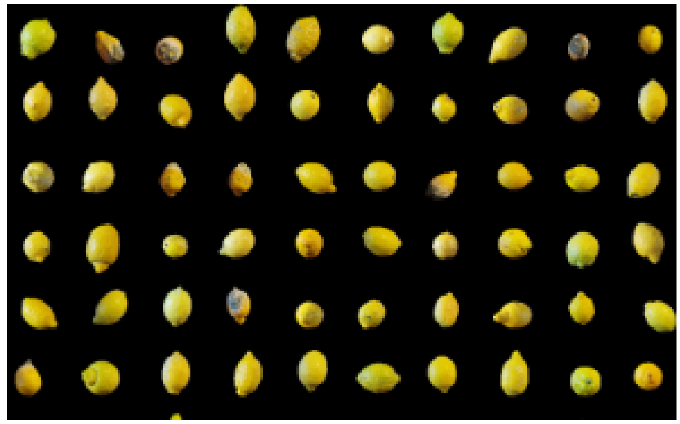

Fig. 1. Lemon Dataset Samples

Next we also utilize Sriram's Fresh Fruit Dataset [1] from Kaggle in the next detector. It comprises of over 13.6k images of Fresh and Rotten - Apples, Bananas, Oranges. Some images sampled form this dataset are shown in Figure 2.

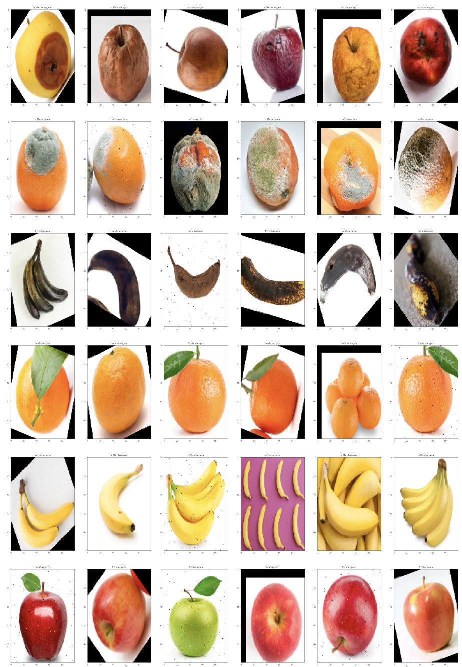

Fig. 2. Fresh Fruits Dataset Samples

We have constructed our Qualiphal model in such a manner that new Data Sources can be added in at anytime and be created into a new detector to improve the quality of the predictions. This modularity helps us quickly and reliably adapt the platform at scale with new sources of data.

B. Model Architectures

1) U-Net: The growing popularity of deep learning and transfer learning have led to various research and studies in image recognition tasks. One such popular CNN based architecture is U-Net which has successfully been used for semantic segmentation in the field of biomedical research[19]. U-Net has a encoder-decoder structure specially designed for the purpose semantic segmentation.Segmentation is nothing but performing classification for every pixel. U-Net involves two two main pathways. The first one involves traditional combinations of a number of convolution layers and max pooling layers. But one issue that arises here is that due to the down sampling and reduction of feature map, the number of features and image size eventually decreases.

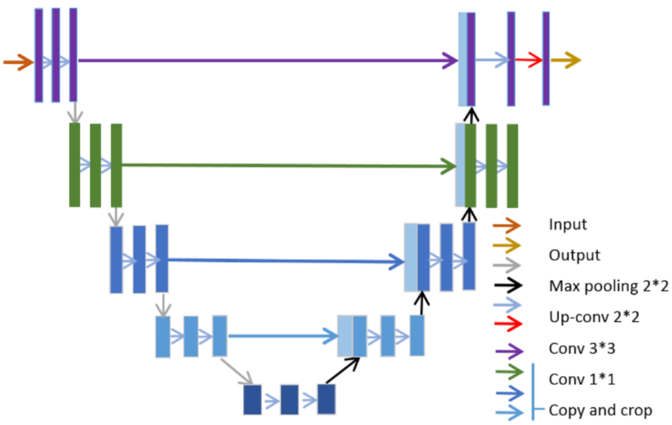

Fig. 3. U-Net Structure

As we aim to perform semantic segmentation, we would require to restore the same size of the image because we need to get classification information for every single pixel and not just the entire image as a whole. Therefore to achieve this U-Net comprises of a second pathway for up sampling which is achieved using transposed convolution which restores the higher dimension of image using back-propagation.

This custom architecture for segmentation that we have implemented is built on top of U-Net architecture uses Efficientnetb3 Figure 4 as backbone. Using such a high performing convolutional network has helped us achieve highly precise localization in the quality inspection process of lemons.

2) FPN: Feature Pyramid Network (FPN) architecture is based on the concept of feature pyramids inside the convolutional neural network (CNN). It has semantically rich and strong features at all levels combining with high resolution via lateral connections. A general architecture for FPN is shown in Figure. 5.

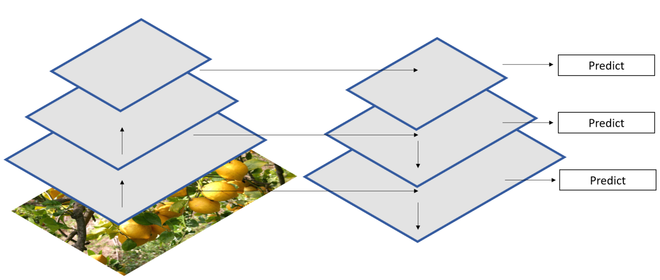

Fig. 5. Feature Pyramid Network

3) Dilation Rate: Dilated convolutions are applied onto an input image to introduce gaps in the feature map. In our work we have made use of dilated convolution for both the proposed networks i.e. U-Net as well as FPN architecture.The dilation rate applied is 8. This has been done to achieve enhanced receptive field, relatively faster computation speed and to cut down the memory usage.

4) Loss: Dice loss (DL) is defined as the negative sum of dice coefficients for each class. Mathematically, it is defined as follows:

$$DL = \sum_{j=1}^{C=M} \frac{2\sum_{i}^{N} p_{ij} g_{ij}}{\sum_{i}^{N} p_{ij}^2 + \sum_{i}^{N} g_{ij}^2} \tag{1}$$

Categorical focal loss (FL) is similar to cross entropy loss that weighs the contribution of each sample to the loss based in the classification error. Mathematically, it is defined as follows:

$$FL = -\sum_{j=1}^{C=M} \sum_{i=1}^{C=N} (1 - t_{ij})^{\gamma} g_{ij} \log (t_{ij}) \tag{2}$$

where M represents the number of classes, p_{ij} represents the binary prediction for i^th voxel j^th class, and g_{ij} represents the binary ground truth for i^th voxel j^th class.

The loss we used for training is defined as the sum of dice loss and categorical focal loss, i.e.

$$L = DL + FL \tag{3}$$

This way we get the advantages of both the loss functions.

C. Application Architecture

1) Training Setup: In this section we discuss about our approach which follow the following steps,

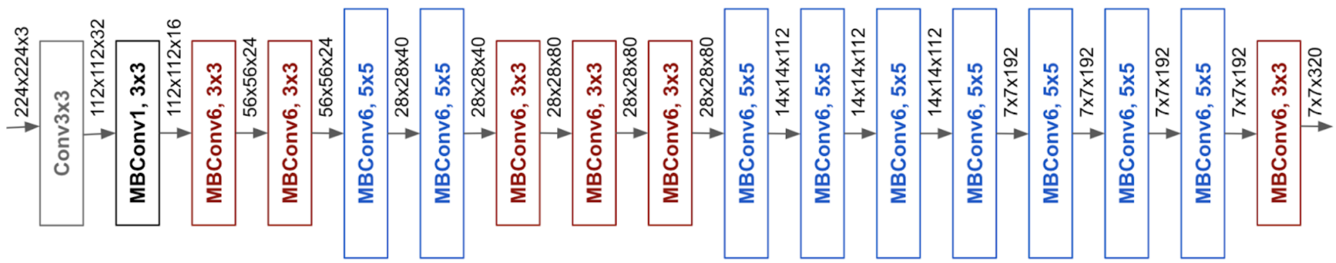

Fig. 4. Efficientnetb3 Structure

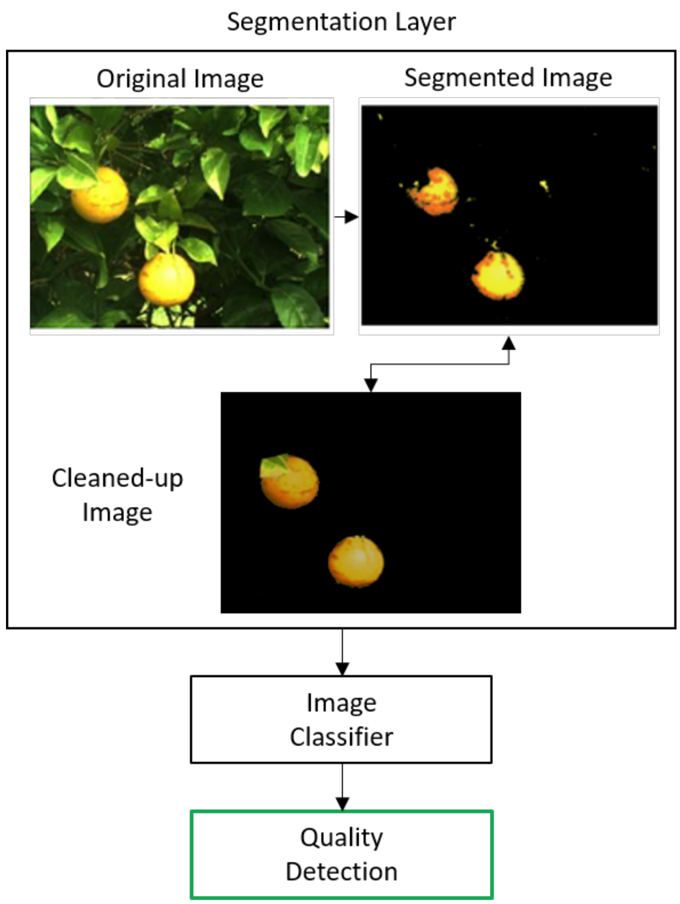

Fig. 6. Qualiphal - Flow Diagram

a Segmentation : Image Segmentation is an important part of the quality detection pipeline as is makes the platform agnostic to environmental conditions and removes the need to have a proper clear background. We account for various conditions like having multiple fruits in the same picture of having a clustered background where it is hard to find the object. The segmentation layer brings the much needed robustness to the platform and makes it applicable in a real world scenario.

b Classification : Image Classification is the second important pillar in making the platform function correctly and reliably. The Classification Layer takes the input from the above segmentation layer and starts classifying them. This is import because here we decide the weights of the detectors, i.e. how much value is their Detection (discussed in the following point) is. This helps us accurately ensemble the outputs weighted by the confidence in the prediction of their class.

c Detection : Quality detection is the central actor in the Qualiphal platform where we are running a multiple detectors in parallel that are trained using a variety of architectures and datasets to widen our prediction base and accurately generalize the ensembled metrics from all the detectors. More detectors can be built and added to improve the quality of the predictions.

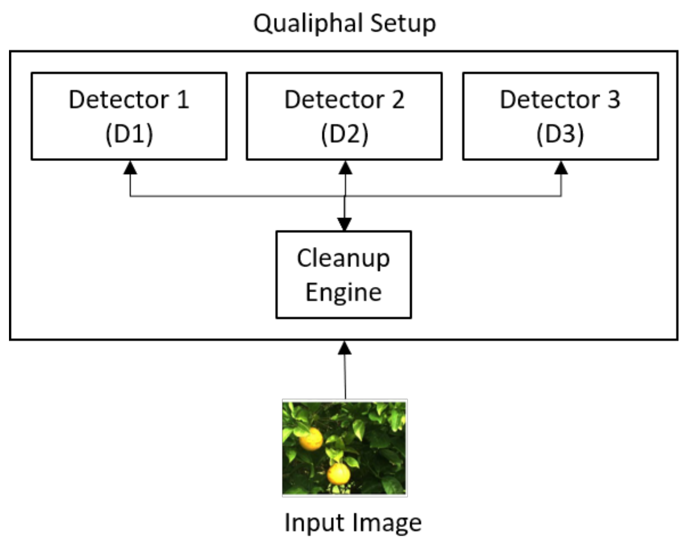

Fig. 7. Qualiphal - Ensemble Setup

2) Ensembled Detectors: The Qualiphal platform was built with the goal of being modular and adaptable. Ensembled detectors help in realizing this goal. From the Classification layer, each object that we detect in the captured image has an associated class list. This class list contains the class of the object and the confidence in the prediction. This setup was chosen as diseases might not be common across the fruits but natural wear and tear can also reduce the quality of the objects. Such a weighted averaging is crucial in the case where if an orange also looks like a lemon we should not overlook the possibility of having a lemon related disease. This approach drastically reduces the false negative and ensures the system can perform reliably even if one of the layers in the network do not.

3) Ease of Accessibility and Scalabilty: Qualiphal is built to be a one-stop-shop for farmers and supply chains where once you are using Qualiphal and it detects that the quality of the fruits are deteriorating, it picks up the geo-location from the device it is being used on and shows a direct map to the nearest drop-off point where the goods can be delivered and consumed immediately. If the fruits are detected diseased, then they alerted for disposal. This ensures the entire supply chin can optimally function without the need for heavy equipment or quality testing labs to check the standard of the food and adapt quickly to environmental or logistical changes.

Qualiphal Application also runs on a serverless platform distributed via a variety of Content Delivery Network (CDN) to handle massive amounts of load. Computationally intensive processes like the Detectors run behind a load balancer that manages and equally distribute the network traffic and help in decided when to scale up or scale down the servers.

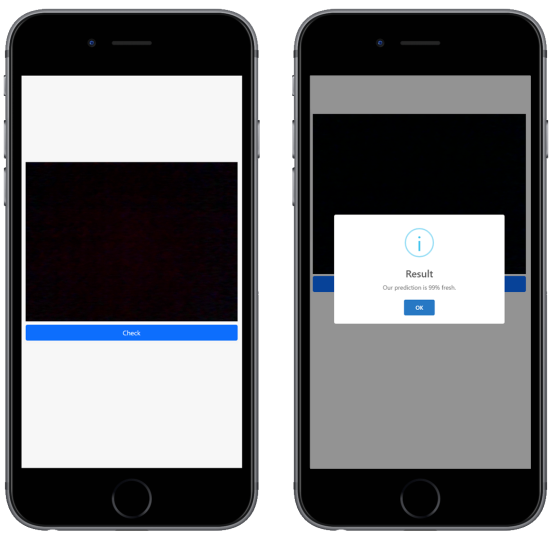

Fig. 8. Application Demo

IV. Experimental Results

	Train			Val		
	Loss	IOU	F1	Loss	IOU	F1
FPN	0.2815	0.7202	0.7205	**0.2173**	**0.7844**	**0.7847**
UNet	0.2810	0.7292	0.7560	0.2244	0.7646	0.7817

In Table IV, we can observe that FPN performs slightly better than UNet on metrics - IoU score(Intersection over Union), F1 and loss defined in above section. Pyramid feature map enriched with lateral connections between bottom-up and top-down pathway gives better segmentation maps when tiny regions are to be segmented from the high resolution images as this is the case in our problem. UNet on the other hand is not able to identify tiny smaller regions but performs excellent on relatively larger segmented portions of fruit.

V. Our Contributions

We implement a custom architecture built on top of Feature Pyramid Network (FPN) which uses the 'Efficient Net' model as backbone which is the main structure of the network. The backbone is obtained from image classification tasks which helps further in image segmentation. We are also comparing the above model with another model where the 'ResNet' model serves as the backbone replacing Efficient Net architecture. We have introduced dilated convolutional layers in both the models which is expected to speed the model by at least 3x.

This helps us not only in improving our metrics like mean Intersection over Union (IoU) but also reduce FLOPS and memory used to train/infer the model. We also do an extensive comparison study between FPN and Unet segmentation models. We evaluated both the models on metrics IoU score as well as focal loss. We conclude the comprehensive study by establishing what type of models work in cases when the segmented area is quite low.

We are creating an open-source framework for COCO-like datasets where it is quite easy to set up the training in just a few lines of code and configuration settings. In less than 10 lines of code, anybody can train their image segmentation model. We provide multiple models, common loss functions and optimizers for instant training setup. Furthermore, it provides extra flexibility to train your own custom models, and loss functions.

This rapid prototyping framework, image segmentation and object detection applications will grow stronger with it. Therefore, there is no need to spend frustrated hours pre-processing and creating data-loaders because everything is available in a few lines of code. It is a high level API for image segmentation tasks easing the application process from data to model saving for production use. We also provide docs for the deployment process to serve end users.

We present an optimized the model using pruning and quantization techniques but also introduced new and different conversion techniques to deploy models on low-end devices. And in doing so we have not lost much of the precision in our prediction results. Therefore, you can capture an image on your mobile device and obtain inference results in seconds. To achieve this a user friendly front end would be created wherein end users can instantly classify whether the food is of bad quality and additionally also understand the causes and reasons for bad quality of the fruit. This way we can process more than ten fruits in seconds. This will hugely benefit in supply chains where not only time is a crucial factor but also checking quality precisely is equally important.

VI. Conclusion

Quality management of fruits and vegetables is very much needed to deliver quality fruits guaranteed to the end consumer. As consumers would mostly pick or reject a particular item by its outer appearance, it is the most important factor while assessing the quality of fruits. The traditional techniques are arduous and labor intensive as they require manual inspection process which is subjective in nature and requires gradual migration to being replaced by more artificially intelligent techniques.

Here we only trained efficient net as the backbone model for both UNet and FPN but experiments with other models such as ResNet, MobileNet and InceptionNet can be carried out. Additionally, experiments combining feature extraction techniques through classical image processing methods like edge detection, and clustering and deep learning methods such as FPN can be feasible way to solve the image segmentation problem for fruit quality.

Acknowledgment

We would like to express our special thanks and gratitude to our Machine learning and Artificial Intelligence professors Dr. Vipul Kumar Mishra and Dr. Apeksha Aggrawal for their able support and guidance to complete this project.

References

[1] Maciej Adamiak. *Fruits - fresh and rotten for classification*. 2018. URL: https://www.kaggle.com/sriramr/fruits-fresh-and-rotten-for-classification/metadata.

[2] Maciej Adamiak. *Lemons quality control dataset*. July 2020. DOI: 10.5281/zenodo.3965568. URL: https://github.com/softwaremill/lemon-dataset.

[3] V. Ashok and D. S. Vinod. "Automatic Fruit Quality Detection System". In: *2019 International Research Journal of Engineering and Technology (IRJET)*. 201. URL: https://www.irjet.net/archives/V6/i6/IRJET-V6I6776.pdf.

[4] V. Ashok and D. S. Vinod. "Automatic quality evaluation of fruits using Probabilistic Neural Network approach". In: *2014 International Conference on Contemporary Computing and Informatics (IC3I)*. 2014, pp. 308–311. DOI: 10.1109/IC3I.2014.7019807.

[5] Irraivan Elamvazuthi et al. "Fruit sorting using fuzzy logic techniques". In: *AIP Conference Proceedings* 1159 (Aug. 2009), pp. 225–230. DOI: 10.1063/1.3223934.

[6] Mandeep Kaur and Reecha Sharma. "Quality Detection of Fruits by Using ANN Technique". In: 10 (July 2015), pp. 35–41. DOI: 10.9790/2834-10423541.

[7] S. Kaur, Akshay Girdhar, and J. Gill. "Computer Vision-Based Tomato Grading and Sorting". In: 2018.

[8] Dalia El Khaled et al. "Fruit and Vegetable Quality Assessment via Dielectric Sensing". In: *Sensors* 15.7 (2015), pp. 15363–15397. ISSN: 1424-8220. DOI: 10.3390/s150715363. URL: https://www.mdpi.com/1424-8220/15/7/15363.

[9] Jonathan Long, Evan Shelhamer, and Trevor Darrell. *Fully Convolutional Networks for Semantic Segmentation*. 2015. arXiv: 1411.4038 [cs.CV].

[10] R. Mottaghi et al. "The Role of Context for Object Detection and Semantic Segmentation in the Wild". In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 891–898. DOI: 10.1109/CVPR.2014.119.

[11] Sapan Naik and Bankim Patel. "Machine Vision based Fruit Classification and Grading - A Review". In: *International Journal of Computer Applications* 170 (July 2017), pp. 22–34. DOI: 10.5120/ijca2017914937.

[12] Zahida Parveen, M. A. Alam, and Hina Shakir. "Assessment of quality of rice grain using optical and image processing technique". In: *2017 International Conference on Communication, Computing and Digital Systems (C-CODE)* (2017), pp. 265–270.

[13] Huy Pham and Byung-Ryong Lee. "An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm". In: *Vietnam Journal of Computer Science* 2 (Feb. 2014), pp. 25–33. DOI: 10.1007/s40595-014-0028-3.

[14] A. Prajwal et al. "Food Quality Detection and Monitoring System". In: *2020 IEEE International Students'*

Conference on Electrical,Electronics and Computer Science (SCEECS). 2020, pp. 1–4. DOI: 10.1109/SCEECS48394.2020.175.

[15] Lennart Reiher, Bastian Lampe, and Lutz Eckstein. *A Sim2Real Deep Learning Approach for the Transformation of Images from Multiple Vehicle-Mounted Cameras to a Semantically Segmented Image in Bird's Eye View*. 2020. arXiv: 2005.04078 [cs.CV].

[16] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. arXiv: 1505.04597 [cs.CV].

[17] Z. Sandoval, F. Prieto, and J. Betancur. "Digital Image Processing for Classification of Coffee Cherries". In: *2010 IEEE Electronics, Robotics and Automotive Mechanics Conference*. 2010, pp. 417–421. DOI: 10.1109/CERMA.2010.54.

[18] Manali R. Satpute and S. M. Jagdale. "Automatic fruit quality inspection system". In: *2016 International Conference on Inventive Computation Technologies (ICICT)* 1 (2016), pp. 1–4.

[19] F. Shi et al. "Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19". In: *IEEE Reviews in Biomedical Engineering* (2020), pp. 1–1. DOI: 10.1109/RBME.2020.2987975.

[20] S. Sivagami et al. "UNet Architecture Based Dental Panoramic Image Segmentation". In: *2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*. 2020, pp. 187–191. DOI: 10.1109/WiSPNET48689.2020.9198370.

[21] Bindu Tiger and T. Verma. "Identification and Classification of Normal and Infected Apples using Neural Network". In: 2013.

[22] Lei Zhou et al. "Application of Deep Learning in Food: A Review". In: *Comprehensive Reviews in Food Science and Food Safety* 18.6 (2019), pp. 1793–1811. DOI: https://doi.org/10.1111/1541-4337.12492. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1541-4337.12492.