Chapter-V

5 Summary and Conclusion

5.1 Summary

This study was initiated with the focus to develop image based deep learning models that can predict the quality of tomatoes from their surface characteristics. In this regard the first three objectives of this study are to develop image-based quality prediction models. The first objective of this study is to apply deep learning-based image processing technique for classifying tomato with physiological attribute-based validation. To accomplish this objective, two different models were developed. To predict the current state of tomatoes as edible and spoilt, a customized CNN model was developed. It was then trained on a self-prepared dataset with 810 images, out of which 572 images are considered for training and 238 images for validation. The model was also trained iteratively with varying epoch and batch sizes to evaluate the model in giving the highest accuracy in classification. The highest accuracy of 99.70% is achieved at epoch 20 and batch size 32. Performance of the developed model was then evaluated by establishing a confusion matrix and Pearson’s correlation. To assist this, sensory evaluation was performed by a group of semi-trained panelists along with physiological analysis. Correlation was established between both using Pearson’s correlation. Pearson’s correlation coefficient R=0.895, precision 100%, recall 87% and accuracy 95% were obtained indicating that model’s prediction ability is at par with human inference. The first model of the first objective is spoilage detection of tomatoes using deep learning.

Secondly, to predict the maturity stage of tomato as mature green, intermediate and advanced, different transfer learning models (VGG, Inception and ResNet) were trained iteratively with varying epoch number and batch size and compared for highest classification accuracy. From the results it was found that VGG 19 performed best at epoch 50 and batch size 32 compared to the other models. On evaluating the performance of the models, Pearson’s correlation coefficient R=0.895, precision 100%, recall 87% and accuracy 95% was obtained. The second model of the first objective is maturity detection of tomatoes using deep learning.

The second objective of this study is to map the pixel-level colour values with physico-chemical properties of tomato using a multi-class classification model.In this objective, an image-based prediction model was developed for estimation of physico-chemical composition of tomatoes based on their surface characteristics. The experiment started with a collection of 120 tomatoes of hybrid variety, analysis of their physico-chemical properties, and image acquisition of the respective tomatoes. Physico-chemical parameters assessed in this study include firmness, color, lycopene content, Titratable Acidity (TA), Total Soluble Solids (TSS), and pH. The parameters are analyzed against their maturity classes, and correlation is established with a statistical tool, Jeffreys's Amazing Statistics Program (JASP). One-way AVONA analysis indicates a statistically significant difference ($p \leq 0.05$) in mean lycopene, TSS, and TA values. Among the parameters, lycopene showed the highest correlation with TSS (R=0.921) and TA (R=-0.858). On the other hand, the least correlated parameter is pH (R=0.393). Lycopene, being the highest correlated chemical parameter, correlation with other physical parameters such as firmness (R=-0.910) and color (R=0.909) is established. The attained physico-chemical results are then mapped against their respective images of three maturity classes. A total of 900 mapped images are then trained in a pre-trained VGG19 model which outperformed the three-class classification with 94% accuracy. The proposed model of the second objective is real-time maturity detection and prediction of physico-chemical properties from tomato image.

The third objective is to develop image processing based multi-class classifier for estimating the shelf-life of tomato. With an objective to estimate the shelf-life of tomatoes, the tomatoes were stored at three different temperatures (5.5±2.2˚C, 18.5±4.9˚C and 29.5±2.1˚C) and its effect on physico-chemical properties of tomatoes was monitored. The degradation kinetics of physico-chemical attributes such as color, firmness, TSS, TA, lycopene content, PWL, PME were estimated using zero, first and second order kinetics. The outcome demonstrates that zero-order model fits best to describe the quality change in tomato upon storage. Furthermore, the reaction rate was assessed using Arrhenius and Eyring model, out of which Arrhenius fits better. Based on the Ea, H and S values, firmness and color are the most critical parameters for shelf-life estimation. Zero-order model combined with Arrhenius model thus predicted shelf-life of tomatoes as 46, 34 and 25 days at 5.5±2.2˚C, 18.5±4.9˚C and 29.5±2.1˚C. Again, to correlate the estimated shelf-life of tomatoes to its surface characteristics, a deep transfer learning model was developed.

To find a real time solution to this, images of respective tomatoes were trained against the outcome of the kinetic model. VGG 19 was successful in predicting the shelf-life of tomatoes with an accuracy of 81%. Hence the developed model of the third objective is to shelf-life prediction of tomatoes from its surface characteristics.

The fourth objective is to deploy the developed model as an application tool for non-destructive on-site assessment of quality and shelf-life of tomato. Recent advancements in deep learning tools such as Tensorflow Lite, have assisted in building a lightweight real-time android-based application for tomato quality inference. Availability of smart phones and its developmental prospects can meet the growing concern of consumers for quality foods from image. Deep learning has significant potential on image identification and hence an image-based application is thus opted. This work is an effort to develop an image-based AI tool for quality inference of tomatoes. To execute the task of application development, the proposed models obtained from the other three objectives were deployed in an android application. Following are the quality inferences (a) prediction of current state of tomatoes as edible or spoilt, mature green, intermediate or advanced (b) prediction of physico-chemical properties (c) shelf-life estimation. The testing and the use of the application, in addition to the high recognition rate, consumes less computation time and is able to make prediction in real-time (<0.67 sec). Thus, this application can be a viable solution in tomato quality inference.

5.2 Conclusion

This study was undertaken to develop deep learning-based models for inference and prediction of quality of tomatoes. As an outcome of this study four different models were developed that can predict quality and shelf-life of tomatoes from image input. The first model developed is a customized CNN model that can be reliable in spoilage detection of tomatoes based on its surface characteristics with an accuracy of 99.70%.In can be concluded that the developed CNN model has the capability in estimating the current state of tomatoes as edible and spoilt can thus be employed in large scale tomato production line. Non-destructive maturity detection of tomatoes using transfer learning techniques was another objective of this study. VGG 19 was successful in automatically classifying tomatoes as-mature green, intermediate and advanced with an accuracy of 97.37% concluding that VGG 19 is a viable image based transfer learning solution to ripening stage classification. VGG 19 is thus considered for the remaining objectives of this study.

In the second objective, an attempt was made to leverage deep learning with food and agriculture. Addressing the current needs of industries, an AI driven tool for automatic internal quality detection of tomatoes was attempted in this objective. VGG19 outperformed the task of mapping the physico-chemical properties of tomatoes against its respective image with an accuracy of 92%. The achieved accuracy shows clear indication of its applicability in industries, and research for real-time maturity detection and prediction of physico-chemical properties from tomato image. Here it can be concluded that the proposed method overcomes various challenges associated with determination of physico-chemical properties through traditional methods.

The fourth model developed can predict the shelf-life of tomatoes based on its surface characteristics with an accuracy of 81%. The proposed model is an automatic solution in predicting the shelf-life on tomatoes from image input. This model is a rapid and accurate solution to minimize tomato losses and will add benefit to the farmer community, in the tomato supply chain, in industries and during storage.

In order to make the developed models available to users, a deep CNN-based android app was developed using TensorFlow lite as the fourth objective of this study. The app developed can infer the quality and its shelf life can be predicted by capturing image using mobile camera in real-time in <0.67 sec from image input. This work is in support of the ongoing digital revolution of India with the advent of AI. Since smart phones have almost reached every nook and corner of India and 5G also has already started rolling out. The developed app will therefore be a low cost, accurate, handy and rapid response to tomato quality and shelf-life.

5.3 Future scope of the thesis

1. The developed image based deep classification models can be validated for a wider variety of fruits and vegetables.
2. The developed app can be extended for assessment of quality and estimation shelf-life for more fruits and vegetables after re-training with more data.