Fruit Maturity Recognition from Agricultural, Market and Automation Perspectives

Koteswar Rao Jerripothula
CSE Department
IIIT-Delhi
New Delhi, India
koteswar@iiitd.ac.in

Sarvesh Kumar Shukla
Cognizant
Pune, India
sarvesh.kumar@cognizant.com

Samyak Jain
CSE Department
Graphic Era
Dehradun, India
samyakatmail@gmail.com

Shudhanshu Singh
Infosys
Pune, India
shudhanshu.singh@infosys.com

IECON 2021 - 47th Annual Conference of the IEEE Industrial Electronics Society | 978-1-6654-3554-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/IECON48115.2021.9589215

Abstract—**Motivated by the potential reduction in the required manual efforts in the fruit industry, this paper attempts to automate fruit maturity recognition. We study the problem from the agricultural, market, and automation perspectives, often taken at different points in the supply chain. Since different maturity states have different visual characteristics, an image classification technology can certainly help here. To develop fruit image classifiers, we need a feature extraction method and a learning algorithm. We use different pre-trained neural networks for effective feature extraction and employ different machine learning algorithms while carrying out bias/variance analysis of the learned models. The analysis helps us select the best ones for each perspective under consideration. We achieve 96%, 94%, and 86% accuracies on our novel dataset named RipeRaw from the agricultural, market, and automation perspectives, respectively.**

Fig. 1. Fruit maturity recognition from three perspectives: (1) agricultural perspective requires specific raw-ripe classifiers; (2) market perspective requires generic raw-ripe classifier; and (iii) automation perspective requires raw-ripe&fruit (multi-class) classifier.

I. Introduction

Automatic fruit maturity recognition [1], [2] is an important problem in the fruit industry, which computer vision can solve. We require this problem to be solved from multiple perspectives such as farms, markets, and robotic assistance. It's required because, while ripened fruits are consumable, the raw ones are not (at least as fruit–some raw fruits such as bananas are consumable as vegetables). Depending upon the consumability of fruits, their packing, delivery planning, variable pricing can be done. The differences in visual characteristics such as color, texture, etc., between raw and ripened fruits make it possible to build a vision-based measurement (VBM) [3], [4], [5], [6] system for automatic[7], [8] maturity recognition. Despite knowing these differences, designing hand-crafted features that can yield very high performance can be very tough, although such features show good promise from the explainability front. In this paper, we study how learned features can help us achieve the required high performance, although at the cost of explainability.

With the rise of deep learning in the last decade, many CNNs were proposed. Since they are trained on millions of images, they can learn rich representations (features) to yield much better performance than the hand-crafted ones. We propose to leverage these learned representations of CNNs to solve the fruit maturity recognition problem with high accuracy.

However, there are several perspectives to this problem. We identify three of them: agriculture [9], [10], market [11], [12] and automation [13], [14] perspectives, as shown in Fig. 1. There is usually a large production of a specific fruit in an agricultural farm, so a dedicated raw-ripe classifier for each fruit is required from the agricultural perspective. On the other hand, there could be a variety of fruits available in the market. The variety also keeps changing, so a generic raw-ripe classifier is required from the market perspective. Both these perspectives so far require building binary classifiers. By automation perspective, we mean a scenario where machines manage the inventory. In such cases, the machines have to perform raw-ripe classification of fruits and recognize them, which makes it a multi-class classification problem.

To cater to the needs of the perspectives discussed, we need to build three types of classifiers. First, to predict if a particular kind of fruit is raw or ripe from the agriculture perspective. Second, to recognize if a given fruit is raw or ripe from the market perspective. Third, to predict both the fruit and its maturity status from the automation perspective. There are many applications to such automated fruit maturity prediction. For example, from the agriculture perspective, we can employ drones to pluck only ripened fruits, separate raw fruits from the ripened ones, and make logistic decisions like where to send which fruits, depending on maturity condition. Likewise, allocating fruits to different refrigeratory resources and maturity-state-based pricing could be applications from the market perspective. Finally, from the automation point of view, fruit-picking and automatic inventory management could be potential applications. In this way, raw-ripe classification has numerous applications in the smooth functioning of the fruit industry.

However, there are some challenges such as (i) lack of publicly available benchmark raw-ripe fruits datasets containing a variety of fruits; (ii) confusion around which feature-extractor and learning algorithm to use to build different classifiers; and (iii) presence of complex backgrounds in natural images.

We handle these challenges in the following manner: (i) We collect ten types of fruit images in raw and ripe states to build a novel dataset named RawRipe dataset. (ii) We search across the existing pre-trained models and learning algorithms for an optimal combination that minimizes bias and variance errors. Then, we build our different final classifiers based on the different optimal combinations we get. Since we need to develop classifiers of three types, and since there are ten varieties of fruits, we build a total of 12 classifiers (10 for agricultural perspective, 1 for market perspective, and 1 for automation perspective). (iii) Since pre-trained models are trained on millions of real-world images, they manage to obtain spatial attention to objects, making backgrounds irrelevant. In this way, by utilizing our new dataset, obtaining optimal combinations of pre-trained networks and learning algorithms, and using CNN's spatial attention capabilities, we could successfully build different classifiers required in the three perspectives.

Our contributions are as follows: (i) Our RawRipe dataset, comprising a total of 1630 images classified into ten fruit categories. (ii) Our methodology helps us identify optimal combinations of pre-trained models and machine learning algorithms while accounting for the bias and variance errors for different classifiers we want to build. (iii) We develop 12 classifiers to solve the raw-ripe fruit classification problem holistically. (iv) Ours is the first attempt to perform raw-ripe fruit classification from multiple (three) perspectives on so many (ten) varieties of fruits.

II. Related Works

Prior works do not consider all the perspectives of raw-ripe classification. Also, most of them only deal with one or two fruits. For example, [15] predicts ripeness of mangoes, and [16] classifies oranges into raw or ripe. Similarly, authors of [17], [18] deal with other fruits like olive and citrus. Most of these solve the problem only from the agricultural perspective, developing a raw-ripe classifier for a particular fruit. Authors of [19], [20], [21] classify a significant number of fruits from the automation perspective but without considering raw-ripe classification. Most of these methods either use hand-crafted features or try to train a feed-forward network from scratch. While the first approach struggles to get high accuracies, the second one requires lots of annotations. Hence, we are first to address the raw-ripe classification problem holistically–that is, we consider multiple perspectives and multiple fruits. Also, we managed to get high accuracies and do not need lots of samples, thanks to our learned features extracted from existing pre-trained networks.

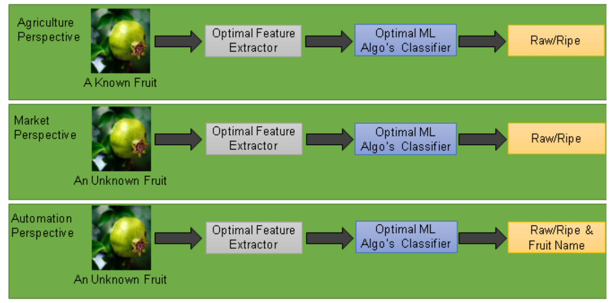

Fig. 2. Our fruit maturity recognition pipelines for different perspectives. All three follow the same process of feature extraction and learning-based classification, but they differ in their input-output combinations.

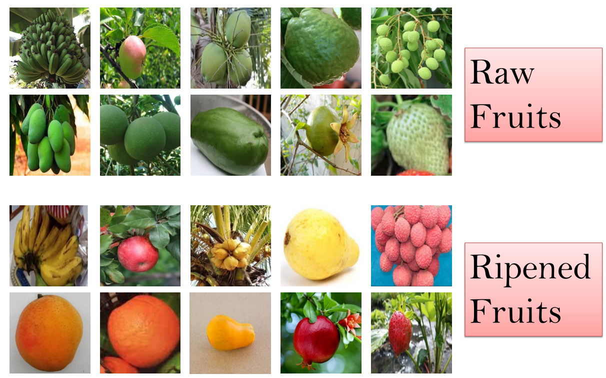

Fig. 3. Sample Images from our RawRipe dataset

III. Proposed Method

We develop a common framework for all three perspectives, as shown in Fig. 2. That is, features are extracted and passed through a learning algorithm's classifier model. However, the input-output pairs of the framework vary across the three perspectives. The framework takes a specific fruit image as input and outputs whether the fruit is raw or ripened in the agricultural perspective. We call the multiple classifiers developed in this perspective specific classifiers. As far as market perspective is concerned, our framework takes any fruit image as input and classifies it as raw or ripe. We call the classifier developed here a generic classifier. In the automation perspective, the framework takes any fruit image as input and predicts its maturity state and the name of the fruit. We call the classifier developed here a multi-class classifier. This section explains how we collect the required data, how we extract features, how we apply machine learning algorithms, and how we select optimal classifiers.

TABLE I
THE DISTRIBUTION OF IMAGES IN OUR RAWRIPE DATASET

	Train		Test		Total
	Raw	Ripe	Raw	Ripe	
Apple	78	81	27	27	213
Banana	63	86	21	29	199
Coconut	54	58	18	20	150
Guava	45	56	16	20	137
Litchi	37	55	13	19	124
Mango	72	46	25	16	159
Orange	48	54	17	18	137
Papaya	51	77	18	25	171
Pomegranate	58	57	20	19	154
Strawberry	65	75	21	25	186
Total	571	645	197	217	1630

A. Data Collection

We collect images for ten types of fruit, both in the raw state and the ripen state: apple, banana, coconut, guava, litchi, mango, orange, pomegranate, papaya, and strawberry. In Fig. 3, we show a few sample images from our dataset. It is a challenging dataset, for the backgrounds of most of them are complicated and can be confused with rawness, as in the case of ripened apple and ripened strawberry images. The details of the data collected for each type are given in Table I. We collect these images by downloading from the web and by capturing them through cameras, as they are pretty common fruits. We manage to obtain images in the range of 137-213 and on an average of 163 for these fruit types. For annotating these images as raw or ripe, we either take web information (if available) or consider the majority voting of three persons. We asked annotators to label the images as raw or ripe by observing the color and texture of the fruits. In this way, we obtain 768 and 862 images with raw and ripe labels, respectively. We name this dataset RawRipe[1]. It comprises 1630 images, split into the training and testing subsets. While we keep 75% of the dataset for training, we use only 25% of it for testing.

B. Perspective-based Data Arrangement

Images in our RawRipe dataset are arranged in three ways: (i) From the agricultural perspective, each fruit requires a separate model to output whether the specific fruit given is raw or ripe. Hence, each row (except the last one, i.e., Total) in Table I becomes a separate dataset to train such models. (ii) From the market perspective, we need only one classifier to predict whether any given fruit is raw or ripe. Hence, the last row, i.e., Total, in Table I becomes the dataset required for this model. (iii) From the automation perspective, we need only one classifier, a multi-class one, to output maturity and the fruit name. It will have 20 labels, i.e., two labels per fruit (e.g., Apple-Raw and Apple-Ripe). By predicting such labels, we know both the maturity and name of the fruit.

[1]The dataset can be downloaded from here: https://sites.google.com/site/koteswarraojerripothula/neatai-servolab/fruit-maturity-recognition

C. Feature Extraction and Machine Learning

Traditionally, machine learning algorithms required hand-crafted features for building effective machine learning models. As a result, only persons with sufficient domain knowledge could design such features. Also, we human beings still don't have a sufficient idea of how exactly our visual cortex works so perfectly in recognizing things, including fruit maturity. Lacking this understanding, it becomes extremely difficult to design excellent features that can help models recognize with perfection, as our brains do. However, the human efforts required to design such features have can significantly reduce through deep learning. The deep neural nets can take the image directly as input and generate rich features as its by-product in addition to the intended predictions when trained on large datasets such as ImageNet. These features can be leveraged to other problems also if they are related to original ones. Numerous pre-trained deep learning models solve the same problem we are trying here, i.e., image classification. Hence, through these pre-trained models, we can easily generate rich features for our problem at hand. We particularly explore Inception v3 [22], VGG-16 [23], VGG-19 [23] and Squeezenet [24] pre-trained deep learning models in our feature extraction step.

Not just features but also the chosen algorithm to learn a classifier also plays a crucial role in the performance. We explore nine algorithms, namely Ada Boost (AB), Decision Tree (DT), k-Nearest Neighbors (kNN), Logistic Regression (LR), Naive Bayes (NB), Artificial Neural Networks (ANN), Random Forests (RF), Stochastic Gradient Descent (SGD), and Support Vectors Machine (SVM) in this paper.

Our next goal is to find an optimal combination of feature extractor and learning algorithm that yields best performance.

Let $\mathcal{F} = \{F_1, \cdots, F_m\}$ be a set of m feature extractors (pre-trained models). Similarly, let $\mathcal{A} = \{A_1, \cdots, A_n\}$ be a set of n algorithms explored. Let (F_i, A_j) represent the combination of feature extractor F_i and algorithm A_j to generate a model. Therefore, a total of $m \times n$ combinations are possible. Let T_{ij} and V_{ij} represent training and cross-validation accuracies, respectively, for i^{th} feature extractor and j^{th} algorithm. We incorporate bias error by comparing the training accuracy with the minimum expected training accuracy, denoted as τ. The higher the bias error, the more simplistic is the predictive model, disregarding important details in the training dataset. If the training accuracy is well above the expected, the bias error will drop. Similarly, we incorporate variance error by comparing the proximity between cross-validation and training accuracies with the maximum expected proximity, denoted as ϵ. The higher the variance error, the more over-fitted is the predictive model. If the proximity between training accuracy and cross-validation accuracy is well below expected, the variance error will drop significantly. Since neither the predictive model should be over-simplistic nor over-fitting, we search for an optimal combination (F_i^*, A_j^*) that maximizes the objective function below:

$$(F_i^*, A_j^*) = \underset{F_i, A_j}{\arg\max}\, \phi\Big(T_{ij}, \tau\Big) + \lambda\psi\Big(|T_{ij} - V_{ij}|, \epsilon\Big) \quad s.t. (F_i, A_j) \in \mathcal{M},\ F_i \in \mathcal{F},\ \text{and}\ A_j \in \mathcal{A} \tag{1}$$

where the first term (ϕ) is the bias term incorporating the bias error issue and the second term (ψ) is the variance term incorporating the variance error issue. There is always a trade-off between bias and variance; therefore, we balance both of these terms using a parameter λ, set as 1.

The bias term, comparing the obtained training accuracy and minimum expected training accuracy, is defined as

$$\phi(T_{ij}, \tau) = T_{ij} - \tau \tag{2}$$

where the more the value, the better the training and the lesser the bias error. Similarly, the variance term that compares the obtained proximity of training accuracy and cross-validation accuracy with the maximum expected proximity is defined as:

$$\psi(|T_{ij} - V_{ij}|, \epsilon) = \epsilon - |T_{ij} - V_{ij}| \tag{3}$$

where the more the value, the better the proximity between training and cross-validation accuracies and the lesser the variance error. Hence, we consider whichever combination (F_i, A_j) maximizes Eqn.(1) over all possible combinations as the optimal one. We conduct an exhaustive search across all possible combinations to find the optimal combination.

IV. Experimental Results

In this section, we discuss the exhaustive experiments conducted. First, we discuss the selection results of our bias-variance analysis to select optimal combinations. Second, we discuss the test results of the models we built. Third, we compare the performance of our models with performances using hand-crafted features.

We conducted all our experiments using Python's Orange library. To have reproducible results without much difficulty, since there are 40 TL models, we use the default settings of its functions. For example, 'k' in the k-fold cross-validation is 5. Similarly, 'k' in the kNN learning algorithm is also 5. We encourage readers to refer to Orange's widget catalog[2] to know more about these default settings.

A. Bias-Variance Analysis Results

From the agricultural perspective, we needed to develop ten specific-type models, one for each fruit. Only optimal combinations are selected based on (1), which is based on the training and cross-validation results we obtain. In Table. II, we provide the cross-validation and training results of only the optimal combinations we obtained. As it can be noted, mostly, the VGG16+LR combination was selected as the optimal combination. While each selected model got 100% training accuracy, the cross-validation accuracy of these models ranged between 94-100%. From the market perspective, we needed to develop a generic model for raw-ripe classification. The cross-validation and training results of all possible combinations are given in Table. III. As it can be seen, the VGG16+LR combination performs the best, again, with training and cross-validation accuracies of 100% and 94.7%, respectively. As far as the automation perspective is concerned, we needed to develop a multi-class model to predict the fruit+maturity labels. The cross-validation and training results of all possible combinations are given in Table. IV. As it can be seen, the VGG16+LR combination performs the best, again, with training and cross-validation accuracies of 100% and 88.2%, respectively. In this way, we could select all the required 12 models for fruit maturity prediction from the three perspectives.

[2]https://orangedatamining.com/widget-catalog/

TABLE II
Chosen optimal combinations for the required specific-type models along with their training and cross-validation accuracies on our RawRipe dataset in relation to agricultural perspective.

Model	**Feature**	**Algorithm**	**Training Accuracy**	**Cross-Validation** Accuracy
Apple	VGG-16	LR	1.000	0.994
Banana	Inception v3	LR	1.000	0.993
Coconut	VGG-16	LR	1.000	1.000
Guava	VGG-16	LR	1.000	0.941
Litchi	VGG-16	LR	1.000	1.000
Mango	Inception v3	LR	1.000	0.958
Orange	VGG-16	LR	1.000	0.990
Papaya	Inception v3	LR	1.000	0.945
Pomegranate	VGG-19	LR	1.000	0.965
Strawberry	VGG-16	ANN	1.000	0.979

TABLE III
Cross-validation/Training Classification Accuracies obtained for different combinations while trying to choose an optimal generic-type model for market perspective. The chosen optimal combination is shown in bold.

	Inception v3	**VGG-16**	**VGG-19**	**SqueezeNet**
AB	0.817 / 1.000	0.840 / 1.000	0.823 / 1.000	0.740 / 1.000
LR	0.938 / 1.000	**0.947 / 1.000**	0.946 / 1.000	0.909 / 1.000
NB	0.822 / 0.843	0.899 / 0.919	0.881 / 0.910	0.747 / 0.754
ANN	0.941 / 1.000	0.945 / 1.000	0.945 / 1.000	0.926 / 1.000
RF	0.884 / 0.998	0.905 / 0.997	0.891 / 0.998	0.822 / 0.992
SGD	0.921 / 1.000	0.937 / 1.000	0.937 / 1.000	0.904 / 0.988
SVM	0.922 / 0.995	0.930 / 0.998	0.934 / 0.992	0.871 / 0.946
DT	0.846 / 0.977	0.852 / 0.972	0.844 / 0.975	0.771 / 0.983
kNN	0.906 / 0.940	0.927 /0.956	0.925 / 0.948	0.846 / 0.905

B. Test Results

The test results of our selected models are given in Table V. We can note that the accuracies are well above 0.9 when we need to perform only raw-ripe classification, i.e., agriculture and market perspective. However, when there is an added responsibility of identifying the fruit, i.e., automation perspective, the performance drops to a 0.85-0.90 range. It is expected because the classifier now has to distinguish between 20 categories, which is more challenging than classifying into two, the case in the other two perspectives. However, all the results are encouraging, for we obtained 96% (overall), 94%, and 86% test accuracy from agriculture, market and automation perspectives, respectively. We give sample qualitative

TABLE IV
CROSS-VALIDATION/TRAINING CLASSIFICATION ACCURACIES OBTAINED FOR DIFFERENT COMBINATIONS WHILE TRYING TO CHOOSE AN OPTIMAL MULTI-CLASS-TYPE MODEL FOR AUTOMATION PERSPECTIVE. THE CHOSEN OPTIMAL COMBINATION IS SHOWN IN BOLD.

	Inception v3	**VGG-16**	**VGG-19**	**SqueezeNet**
AB	0.544 / 1.000	0.529 / 1.000	0.516 / 1.000	0.363 / 1.000
LR	0.879 / 1.000	**0.882 / 1.000**	0.864 / 1.000	0.770 / 1.000
NB	0.757 / 0.874	0.756 / 0.892	0.747 / 0.891	0.540 / 0.698
ANN	0.871 / 1.000	0.868 / 1.000	0.850 / 1.000	0.764 / 1.000
RF	0.662 / 0.989	0.697 / 0.995	0.661 /0.995	0.506 / 0.984
SGD	0.822 / 1.000	0.809 / 1.000	0.814 / 1.000	0.678 / 0.961
SVM	0.842 / 0.985	0.805 / 0.989	0.793 / 0.986	0.743 / 0.949
DT	0.567 / 0.930	0.538 / 0.920	0.501 /0.928	0.377 / 0.899
kNN	0.766 / 0.852	0.749 / 0.845	0.735 /0.842	0.607 / 0.767

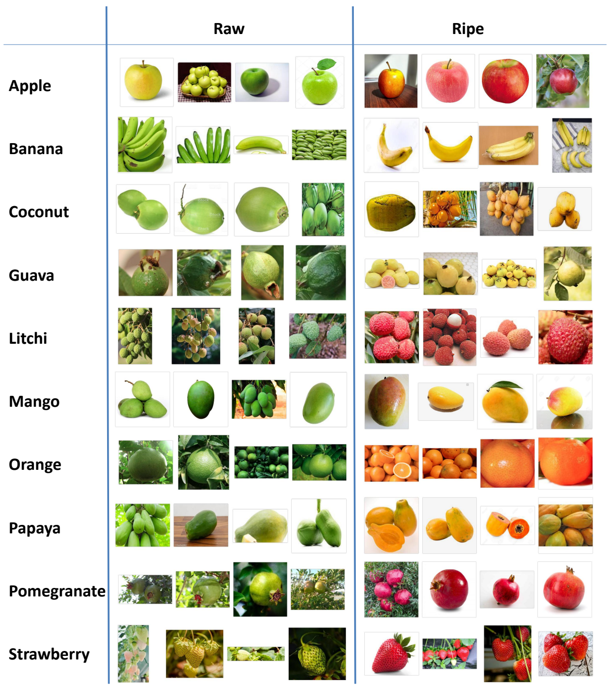

Fig. 4. Sample results from the agricultural perspective.

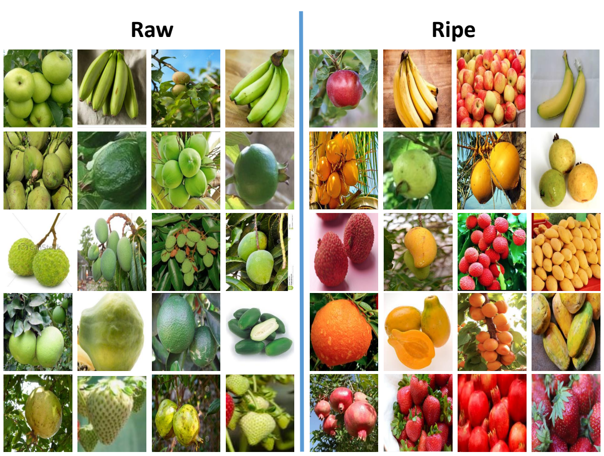

Fig. 5. Sample results from the market perspective.

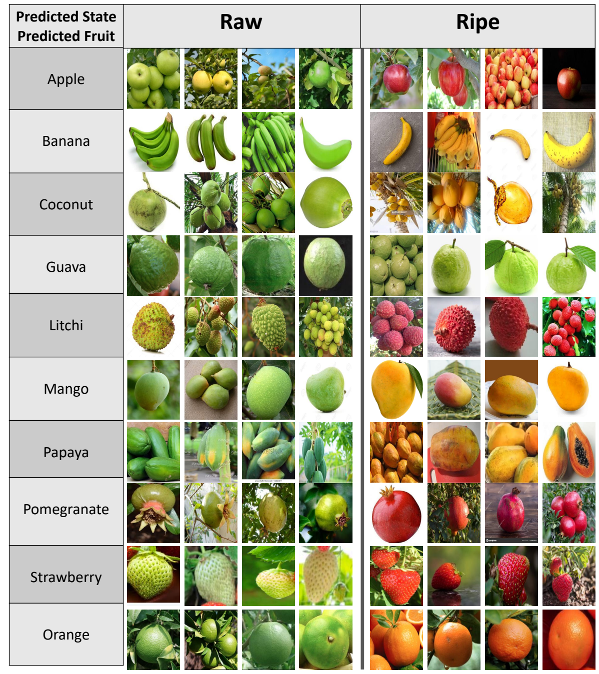

Fig. 6. Sample results from the automation perspective.

TABLE V
TESTING CLASSIFICATION ACCURACIES OF CHOSEN PREDICTIVE MODELING APPROACHES FOR OUR THREE TYPES OF MODELS

Model	**Feature**	**Algorithm**	**Test Accuracy**
Apple	VGG-16	LR	1.000
Banana	Inception v3	LR	0.960
Coconut	VGG-16	LR	1.000
Guava	VGG-16	LR	0.943
Litchi	VGG-16	LR	0.969
Mango	Inception v3	LR	0.927
Orange	VGG-16	LR	1.000
Papaya	Inception v3	LR	0.907
Pomegranate	VGG-19	LR	1.000
Strawberry	VGG-16	ANN	0.913
Generic	VGG-16	LR	0.944
Multi-class	VGG-16	LR	0.862

results from agricultural, market and automation perspectives in Fig. 4, Fig. 5 and Fig. 6, respectively. We can see that our classifiers can classify the fruit images quite well.

C. Comparisons

We compare our generic and multi-class models with those developed using hand-crafted features like color Histogram, GIST, and their concatenation, $\oplus$, in Table VI and Table VII, respectively. We can see that our reliance on learned features extracted from pre-trained networks yields superior results. Although hand-crafted features yielded decent results for the 2-class problem (market), they perform poorly for the challenging 20-class problem (automation). This is because those hand-crafted features can only capture color and texture required for raw-ripe classification, but classifying fruits based

TABLE VI
COMPARISON WITH HAND-CRAFTED FEATURES WHILE CONSIDERING THE MARKET PERSPECTIVE

Model	Training	Cross-validation	Testing
GIST+LR	0.650	0.609	0.640
HIST+LR	0.819	0.806	0.853
GIST$\oplus$HIST+LR	0.838	0.797	0.855
Ours (VGG16+LR)	1.000	0.947	0.944

TABLE VII
COMPARISON WITH HAND-CRAFTED FEATURES WHILE CONSIDERING THE AUTOMATION PERSPECTIVE

Model	Training	Cross-validation	Testing
GIST+LR	0.347	0.280	0.271
HIST+LR	0.262	0.217	0.258
GIST$\oplus$HIST+LR	0.501	0.394	0.408
Ours (VGG16+LR)	1.000	0.882	0.862

on those hand-crafted ones is expected to be difficult because that requires capturing semantic information, which needs learning. Hence, utilizing learned features becomes crucial for effective fruit maturity recognition.

From the Table VII, we can also note that there is a significant improvement when we combine the two handcrafted-features compared to their individual performances. Although combination helps, the performance is way below compared to what we get through learned features because even efficient combination requires learning.

CONCLUSION

We considered the raw-ripe classification problem for fruits from three perspectives: agricultural, market, and automation. For this purpose, we created a novel dataset named RawRipe, covering natural raw and ripened images of as many as ten fruits. We then performed bias/variance analysis to find optimal combination of pre-trained models (for feature extraction) and learning algorithms to select our models. We obtained 96% (overall), 94%, and 86% test accuracy from agriculture, market and automation perspectives, respectively.

ACKNOWLEDGEMENTS

This work was supported in part by the Infosys Centre for Artificial Intelligence, IIIT-Delhi, and in part by IIIT-Delhi PDA Grant.

REFERENCES

[1] A. Bhargava and A. Bansal, "Quality evaluation of mono & bi-colored apples with computer vision and multispectral imaging," *Multimedia Tools and Applications*, pp. 1–18, 2020.

[2] S. Wan and S. Goudos, "Faster r-cnn for multi-class fruit detection using a robotic vision system," *Computer Networks*, vol. 168, p. 107036, 2020.

[3] D. S. Leme, S. A. da Silva, B. H. G. Barbosa, F. M. Borém, and R. G. F. A. Pereira, "Recognition of coffee roasting degree using a computer vision system," *Computers and electronics in agriculture*, vol. 156, pp. 312–317, 2019.

[4] K. Kärhä, S. Nurmela, H. Karvonen, V.-P. Kivinen, T. Melkas, and M. Nieminen, "Estimating the accuracy and time consumption of a mobile machine vision application in measuring timber stacks," *Computers and electronics in agriculture*, vol. 158, pp. 167–182, 2019.

[5] S. A. Ansari, K. R. Jerripothula, P. Nagpal, and A. Mittal, "Eye-focused detection of bell's palsy in videos," *Proceedings of the Canadian Conference on Artificial Intelligence*, 6 2021, https://caiac.pubpub.org/pub/lk3jdnfp. [Online]. Available: https://caiac.pubpub.org/pub/lk3jdnfp

[6] D. Goyal, K. Rao Jerripothula, and A. Mittal, "Detection of gait abnormalities caused by neurological disorders," in *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, 2020, pp. 1–6.

[7] K. R. Jerripothula, J. Cai, F. Meng, and J. Yuan, "Automatic image co-segmentation using geometric mean saliency," in *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 3277–3281.

[8] K. Rao Jerripothula and P. Mukherjee, "Asoc: Adaptive self-aware object co-localization," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*, 2021, pp. 1–6.

[9] S. AlZu'bi, B. Hawashin, M. Mujahed, Y. Jararweh, and B. B. Gupta, "An efficient employment of internet of multimedia things in smart and future agriculture," *Multimedia Tools and Applications*, vol. 78, no. 20, pp. 29 581–29 605, 2019.

[10] N. Deepa and K. Ganesan, "Decision-making tool for crop selection for agriculture development," *Neural Computing and Applications*, vol. 31, no. 4, pp. 1215–1225, 2019.

[11] B.-H. Zhou and F. Tan, "A self-adaptive estimation of distribution algorithm with differential evolution strategy for supermarket location problem," *Neural Computing and Applications*, pp. 1–14, 2019.

[12] F. T. García, L. J. G. Villalba, A. L. S. Orozco, and T.-H. Kim, "A comparison of learning methods over raw data: forecasting cab services market share in new york city," *Multimedia Tools and Applications*, vol. 78, no. 21, pp. 29 783–29 804, 2019.

[13] D. Popa, F. Pop, C. Serbanescu, and A. Castiglione, "Deep learning model for home automation and energy reduction in a smart home environment platform," *Neural Computing and Applications*, vol. 31, no. 5, pp. 1317–1337, 2019.

[14] P. P. Roy, P. Kumar, and V. Chang, "A hybrid classifier combination for home automation using eeg signals," *Neural Computing and Applications*, pp. 1–13, 2020.

[15] D. Sahu and R. M. Potdar, "Defect identification and maturity detection of mango fruits using image analysis," *American Journal of Artificial Intelligence.*, vol. 1, no. 1, pp. 5–14, 2017.

[16] B. Kanimozhi and R. Malliga, "Classification of ripe or unripe orange fruits using the color coding technique," *Asian Journal of Applied Science and Technology (AJAST) Volume*, vol. 1, pp. 43–47, 2017.

[17] D. Li, Y. Liu, and Y. Chen, Eds., *Computer and Computing Technologies in Agriculture IV - 4th IFIP TC 12 Conference, CCTA 2010, Nanchang, China, October 22-25, 2010, Selected Papers, Part I*, ser. IFIP Advances in Information and Communication Technology, vol. 344. Springer, 2011. [Online]. Available: https://doi.org/10.1007/978-3-642-18333-1

[18] R. C. Wilson, E. R. Hancock, A. G. Bors, and W. A. P. Smith, Eds., *Computer Analysis of Images and Patterns - 15th International Conference, CAIP 2013, York, UK, August 27-29, 2013, Proceedings, Part I*, ser. Lecture Notes in Computer Science, vol. 8047. Springer, 2013. [Online]. Available: https://doi.org/10.1007/978-3-642-40261-6

[19] A. Patino-Saucedo, H. Rostro-Gonzalez, and J. Conradt, "Tropical fruits classification using an alexnet-type convolutional neural network and image augmentation," in *International Conference on Neural Information Processing*. Springer, 2018, pp. 371–379.

[20] Y. Zhang, S. Wang, G. Ji, and P. Phillips, "Fruit classification using computer vision and feedforward neural network," *Journal of Food Engineering*, vol. 143, pp. 167–177, 2014.

[21] Y. Zhang, P. Phillips, S. Wang, G. Ji, J. Yang, and J. Wu, "Fruit classification by biogeography-based optimization and feedforward neural network," *Expert Systems*, vol. 33, no. 3, pp. 239–253, 2016.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations*, 2015.

[24] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and $<$1mb model size," *CoRR*, vol. abs/1602.07360, 2016.