Knowledge on Visual representation of binary classification: Apple Vs Avocado

Dr.G.Ayyappan[1], Dr.R.Venkatesh Babu[2], Dr.K.Alice[3], Dr.A.Kumaravel[4]

Associate Professor, Department of Information Technology, Prince Shri Venkateshwara Padmavathy Engineering College, Chennai[1]

Professor, Pro-Vice Chancellor, Bharath Institute of Higher Education and Research, Chennai[2]

Assistant Professor, Department of Computing Technologies, SRM Institute of Science and Technology, SRMIST KTR Campus, Chennai[3]

Professor, Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai[4]

Email Id: ayyappangmca@gmail.com[1], rvbaboo76@gmail.com[2], alicesureshme@gmail.com[3], drkumaravel@gmail.com[4]

Abstract:

Visual cue analysis for fruit categorization and sorting automates the visual inspection and packing process in agricultural applications. Similarities in fruit color and shape challenge automated multi-class sorting systems. Deep neural networks require extensive training datasets to achieve high accuracy, which is not available for agricultural items, notably fruits and vegetables. This work governs that the Auto Color Correlogram (ACCF) and Binary Pyramid Patter filter (BPPF) model using Bayes Net (BN) learning model. The ACCF with BN gives accuracy 95%, BPPF with BN gives accuracy 55%. The ACCF with BN gives precision 0.95, BPPF with BN gives precision 0.55. The ACCF with BN gives recall 0.95, BPPF with Bayes Net gives recall 0.55. The ACCF with BN gives F-Measure 0.95, BPPF with BN gives F-Measure 0.55.The ACCF with BN gives MCC 0.91, BPPF with Bayes Net gives MCC 0.1.The ACCF with Bayes Net gives ROC 0.94, BPPF with Bayes Net gives ROC 0.63. The ACCF with Bayes Net gives PRC 0.93, BPPF with BN gives PRC 0.61. The ACCF with BN gives Kappa 0.9, Binary Patter Pyramid Filter with BN gives Kappa 0.1.The ACCF with BN gives mean absolute error 0.05, BPPF with BN gives MAE 0.43. The ACCF with BN gives root mean squared error 0.22, BPPF with BN gives RMSE 0.53. The ACCF with BN gives relative absolute error 10%, BPPF with BN gives RAE 85.46%. The ACCF with BN gives root relative squared error 44.72%, BPPF with BN gives RRSE 107.83%. The ACCF with BN model takes 0.49 seconds for creating its model, BPPF with Bayes Net takes 0.05 seconds for creating its model. The ACCF with BN performs well compared with BPPFF with BN model.

Keywords: accuracy, avocado, Bayes Net, apple, Binary Pyramid Pattern

I Introduction

Agricultural automation includes automated fruit classification and sorting. Using robotic platforms, automated fruit classification systems can be employed during harvest for fruit identification and sorting. Also used in the packaging industry for post-harvest quality assessment, fruit harvesting, and pricing identification in supermarkets for quick billing. To improve fruit quality and productivity, manual fruit picking and sorting has been replaced with automated systems using machine vision and machine learning. Since classification accuracy depends on feature quality, many studies have compared feature sets in different datasets. Many scholars have attempted various ways of fruit classification.[1,5-6]. The research retrieved fruit color, size, shape, and texture. Decision Trees and bagging classifier were compared.[7] . Their system was 95.3%. SVM classifies rotten fruit. A six-layered CNN and VGG-16 on two datasets. The VGG-16 model obtained 99.75% accuracy on the supermarket produce dataset and 96.75% on the self-collected dataset.[11]. Authors classified fruit for supermarkets.[12] Six apple cultivars were classified using a 9-layer deep neural network. The authors classify fruits by color and texture.[13-16] The system identified 15 fruit classes using two features. Their dataset consists of internet-sourced white-background fruit photos.[17] .The KNN classifier's accuracy was 81.9%. The system compared machine learning strategies for fruit classification.[18] Apple, banana, orange, pear, watermelon, and mango were considered.[19] They found areas of interest using thresholding and morphological processing.[20] An algorithm to categorize four fruit kinds and detect rotten fruit. The KNN classifier identifies fruit by color and feel.[21] . Area, color, centroid, zone, perimeter, size, roundness were extracted. KNN, SVM, Naive Bayes, random forest, and neural networks were employed. The SVM classifier had the highest accuracy, at 91.67%. [22]. They used HSV thresholding to extract ROI. After using a three-level discrete wavelet transform, they extracted hue, saturation, and brightness information. They identified 10 fruit classes using an SVM

classifier.[23,24]The system governed the creation of a CNN-based classifier that recognized 25 fruit classes with 100% test and 99.79% train accuracy[25]. They used zoning, character edge descriptors, and DFT. Classifiers included MLP and KNN. It reported 97.5% accuracy. Current work aims to create a durable, efficient, and repeatable method for identifying fruits using handcrafted features.[26] Handcrafted features include edges, corners, histograms, etc. that are extracted manually from image data.[27] Deep neural networks that use automatically learned features require a lot of training data per class to produce decent classification results, which is frequently unavailable for agricultural items like fruits and vegetables.[28]Many studies using handmade features and typical machine learning techniques for fruit categorization lack accuracy and repeatability. Handcrafted features should improve accuracy for datasets with fewer data points per class.[29] Another goal is to compare supervised learning algorithms to identify the best fruit classifier using custom characteristics.[30] By extracting appropriate characteristics and employing an efficient classifier, automated fruit categorization and sorting will be improved. We present a unique combination of Hue, CSIFT, Discrete Wavelet Transform, and Haralick features that yields good accuracy in fruit categorization. Authors developed a grape recognition system for fruit-picking robots[31].

The rest of the paper: Section 2 explains the process. Section 3 provides the results. Section 4 discusses results-derived observations. Section 5 concludes.

II Data Collection and Research Methods

The data set has been collected from the kaggle dataset, which is namely the Fruit classification (10 Class). It consists of 3374 images and is a collection of the following images 1. Apple, 2. Kiwi, 3. Banana, 4. Cherry, 5. Orange, 6. Mango, 7. Avocado, 8. Pineapple, 9.Strawberries, and 10. Watermelon. In this experiment, we considered only apples and avocados for binary image classification. The collected 60 images have been taken for the total population, which was implemented into 90% of training data and 10% of test data in the Weka data mining tool.
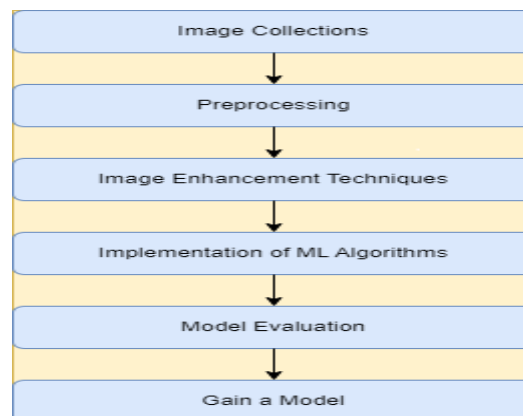

Figure 1: Proposed System

Methodology:

- Preprocessing the data (randomly chosen data has already been preprocessed various dimensions into 256*256 dimension of JPEG images).
- Prepare image data that has been selected in CSV format.
- Insert the dataset into Weka.
- Tuning the required parameters in Weka.
- Extract attributes by applying the Attribute Color Correlogram Filter technique.
- Apply Bayes' Net Classification.
- Extract attributes by applying the Binary Patterns Pyramid Filter technique.
- Apply Bayes' Net Classification.

The above flow process has been implemented to discover a new model.

IV Results and Discussions

This section focuses on the outcome of this research work by using several parameter tuning and selected methodology.

In ACCF+BN technique, If the maximize cost 30, random 30 and gain 0, this model gives classification accuracy is 50%, which means TP 50%(30 images),TN 0%(0 images),FP 50%(30 images), and FN 0%(0 images); If the cost 3, random 30 and gain 27, this model gives overall classification accuracy is 95%, which means TP 45%(27 images), TN 5%(3 images), FP 0%(0 images), and FN 50%(30 images); When maximize cost 30, random 30 and gain 0, this model gives classification accuracy is 50%, which means TP 50%(30 images),TN 0%(0 images),FP 50%(30 images),and FN 0%(0 images); If the cost 3, random 30 and gain 27, this model gives overall classification accuracy is 95%,which means TP 50%(30 images), TN 0%(0 images), FP 5%(3 images), and FN 45%(27 images); If the benefit 30, random 30 and gain 0, this model gives overall classification accuracy is 95%,which means TP 50%(30 images), TN 0%(0 images), FP 50%(30 images), and FN 0%(0 images); If the benefit 3, random 30 and gain -27, this model gives overall classification accuracy is 95%,which means TP 45%(27 images), TN 5%(3 images), FP 0%(0 images), and FN 50%(30 images); If the minimize benefit 30, random 30 and gain 0, this model gives classification accuracy is 50%, which means TP 50%(30 images),TN 0%(0 images),FP 50%(30 images), and FN 0%(0 images); The minimize benefit of avocado class when apply ACCF+BN technique. If the benefit 3, random 30 and gain -27, this model gives overall classification accuracy is 95%,which means TP 50%(30 images), TN 0%(0 images), FP 5%(3 images), and FN 45%(27 images);The maximize cost of apple class when apply BPPFF+BN technique. If the maximize cost 31, random 30 and gain -1, this model gives classification accuracy is 48.33%, which means TP 48.33%(29 images),TN 1.67%(1 images),FP 50%(30 images), and FN 0%(0 images).

In BPPF+BN technique, If the cost 21, random 30 and gain 9, this model gives overall classification accuracy is 65%, which means TP 438.33%(23 images), TN 11.67%(7 images), FP 23.33%(14 images), and FN 26.67%(16 images); If the benefit 31, random 30 and gain 1, this model gives overall classification accuracy is 48.33%,which means TP 48.33%(29 images), TN 1.67%(1 images), FP 50%(30 images), and FN 0%(0 images); If the benefit 21, random 30 and gain -9, this model gives overall classification accuracy is 65%,which means TP 38.33%(23 images), TN 11.67%(7 images), FP 23.33%(14 images), and FN 26.67%(16 images); When maximize cost 31, random 30 and gain -1, this model gives classification accuracy is 48.33%, which means TP 1.67%(1 image),TN 48.33%(29 images),FP 3.33%(2 images),and FN 46.67%(28 images); If the cost 21, random 30 and gain 9, this model gives overall classification accuracy is 65%,which means TP 26.67%(16 images), TN 23.33%(14 images), FP 11.67%(7 images), and FN 38.33%(23 images); If the minimize benefit 31, random 30 and gain 1, this model gives classification accuracy is 48.33%, which means TP 1.67%(1 images),TN 48.33%(29 images),FP 3.33%(2 images), and FN 46.67%(28 images); If the benefit 21, random 30 and gain -9, this model gives overall classification accuracy is 65%,which means TP 26.67%(16 images), TN 23.33%(14 images), FP 11.67%(7 images), and FN 38.33%(23 images).

Table 1: Log score of ACCF+BN and BPPF+BN

S.No	Model	Local Score	Values
1	ACCF+BN	Bayes	-2104.7561574090546
2		Bayesian Database (BDeu)	-2237.03449138485
3		Minimum Description Length	-2314.339921426314
4		ENTROPY	-1997.0282178541008
5		Akaike Information Criterion (AIC)	-2152.0282178541
6	BPPF+BN	Bayes	-194.33150307960796
7		Bayesian Database (BDeu)	-204.638890572707254
8		Minimum Description Length	-212.4461682066894
9		ENTROPY	-185.83292855224576
10		Akaike Information Criterion (AIC)	-198.83292855224576

The above table shows the ACCF with BN produces several local scores like Bayes, BDeu, MDL, Entropy and AIC values.
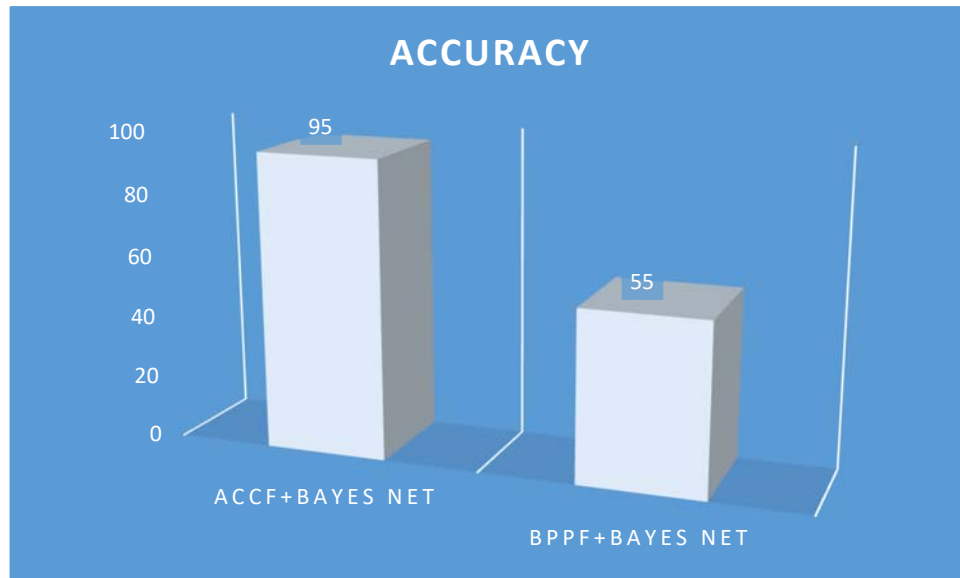
Figure 2: Representation of Accuracy

The above diagram 39 shows that the Auto Color Correlogram Filter with Bayes Net gives accuracy 95%, Binary Patter Pyramid Filter with Bayes Net gives accuracy 55%. ACCF with BN performs well compared with BPPF with BN model.
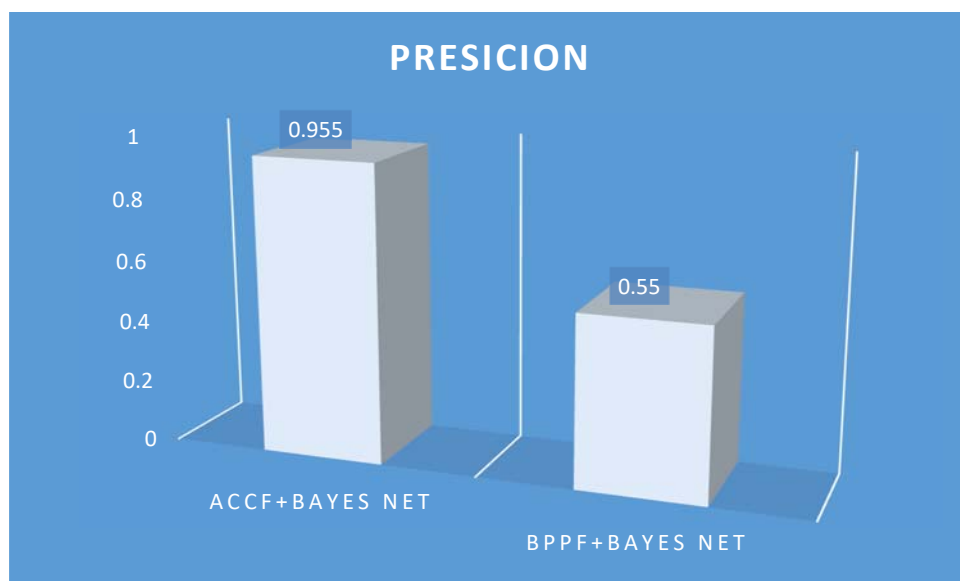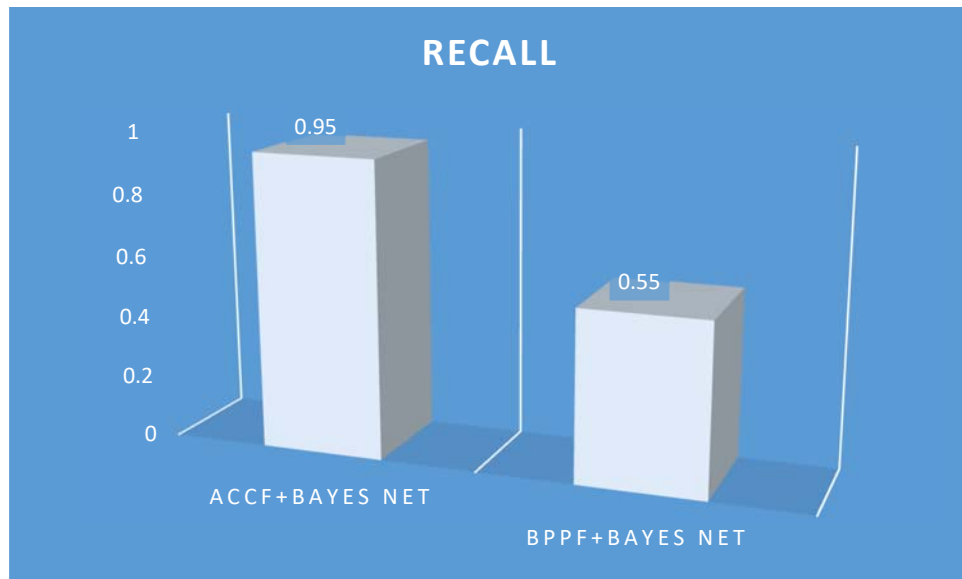

Figure 3: Representation of Precision

The above diagram 40 shows that the Auto Color Correlogram Filter with Bayes Net gives precision 0.95, Binary Patter Pyramid Filter with Bayes Net gives precision 0.55. ACCF with BN performs well compared with BPPF with BN model.

Figure 4: Representation of Recall

The above diagram 41 shows that the Auto Color Correlogram Filter with Bayes Net gives recall 0.95, Binary Patter Pyramid Filter with Bayes Net gives recall 0.55. ACCF with BN performs well compared with BPPF with BN model.
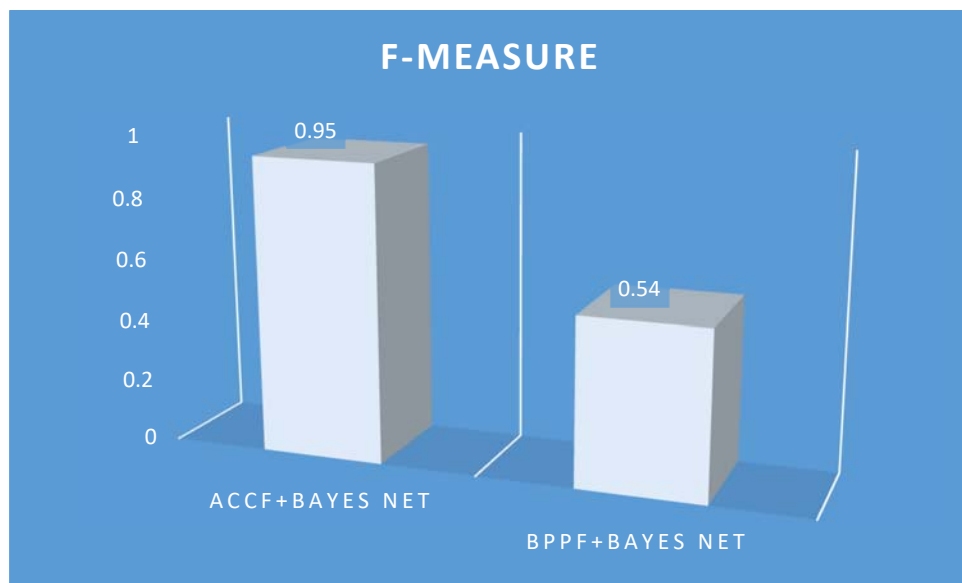

Figure 5: Representation of F-Measure

The above diagram 42 shows that the Auto Color Correlogram Filter with Bayes Net gives F-Measure 0.95, Binary Patter Pyramid Filter with Bayes Net gives F-Measure 0.55. ACCF with BN performs well compared with BPPF with BN model.
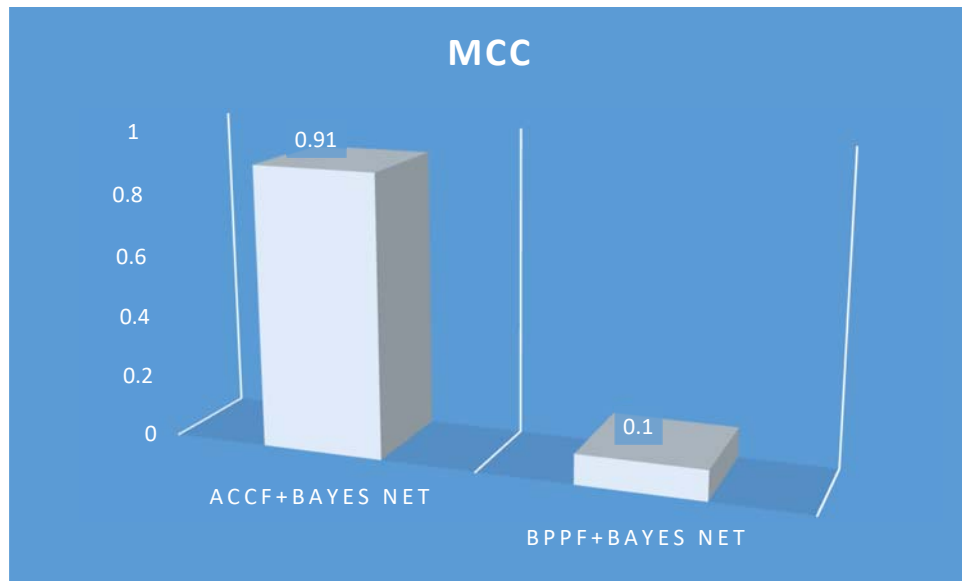
Figure 6: Representation of MCC

The above diagram 43 shows that the Auto Color Correlogram Filter with Bayes Net gives MCC 0.91, Binary Patter Pyramid Filter with Bayes Net gives MCC 0.1. ACCF with BN performs well compared with BPPF with BN model.
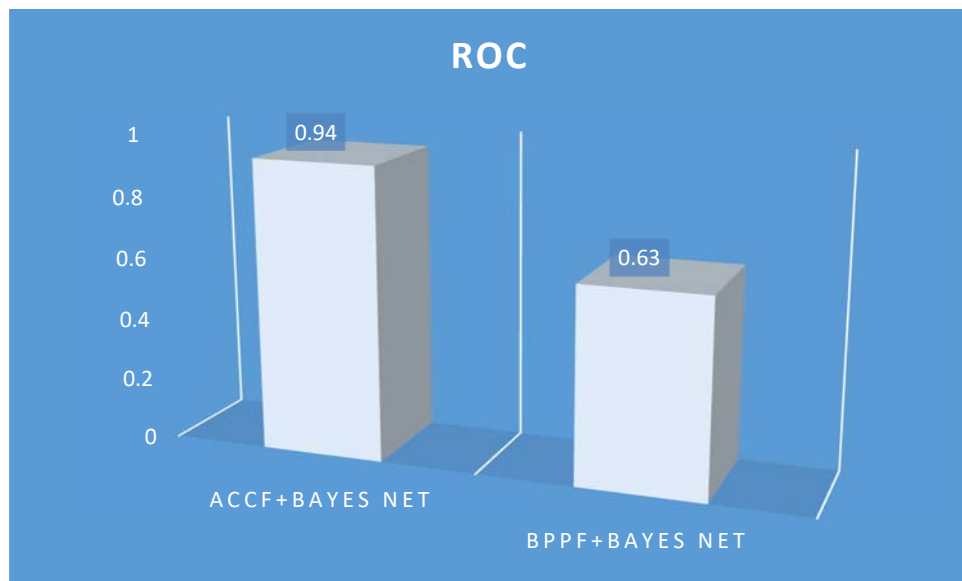

Figure 7: Representation of ROC

The above diagram 44 shows that the Auto Color Correlogram Filter with Bayes Net gives ROC 0.94, Binary Patter Pyramid Filter with Bayes Net gives ROC 0.63. ACCF with BN performs well compared with BPPF with BN model.
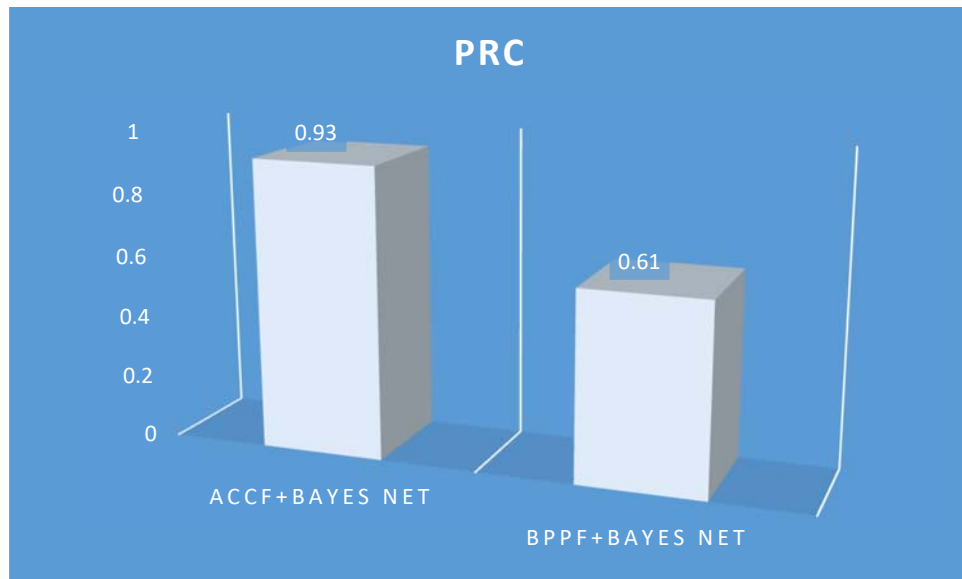
Dr.G.Ayyappan et al. / Indian Journal of Computer Science and Engineering (IJCSE)

Figure 8: Representation of PRC

The above diagram 45 shows that the Auto Color Correlogram Filter with Bayes Net gives PRC 0.93, Binary Patter Pyramid Filter with Bayes Net gives PRC 0.61. ACCF with BN performs well compared with BPPF with BN model.
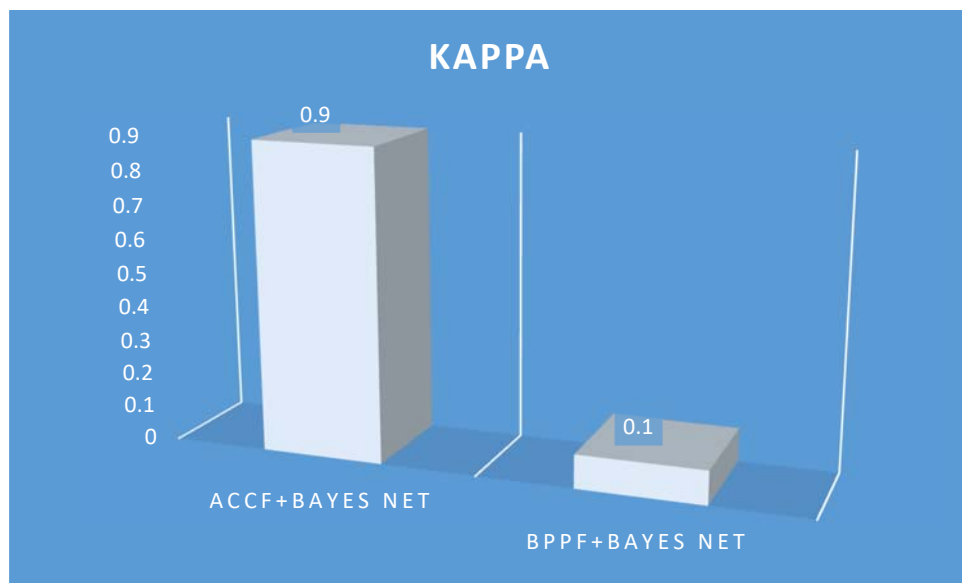

Figure 9: Representation of Kappa

The above diagram 46 shows that the Auto Color Correlogram Filter with Bayes Net gives Kappa 0.9, Binary Patter Pyramid Filter with Bayes Net gives Kappa 0.1. ACCF with BN performs well compared with BPPF with BN model.
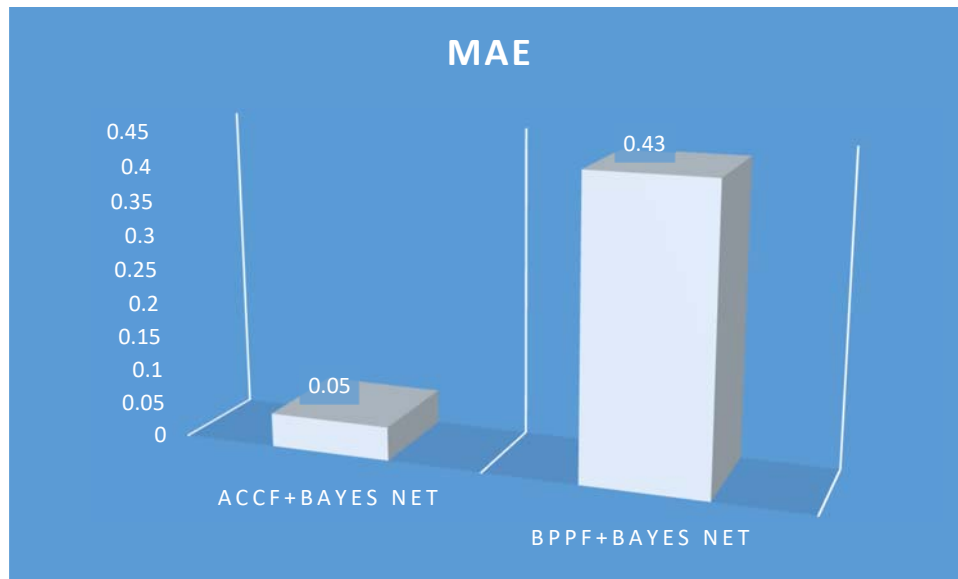
Figure 10: Representation of MAE

The above diagram 47 shows that the Auto Color Correlogram Filter with Bayes Net gives mean absolute error 0.05, Binary Patter Pyramid Filter with Bayes Net gives MAE 0.43. ACCF with BN performs well compared with BPPF with BN model.
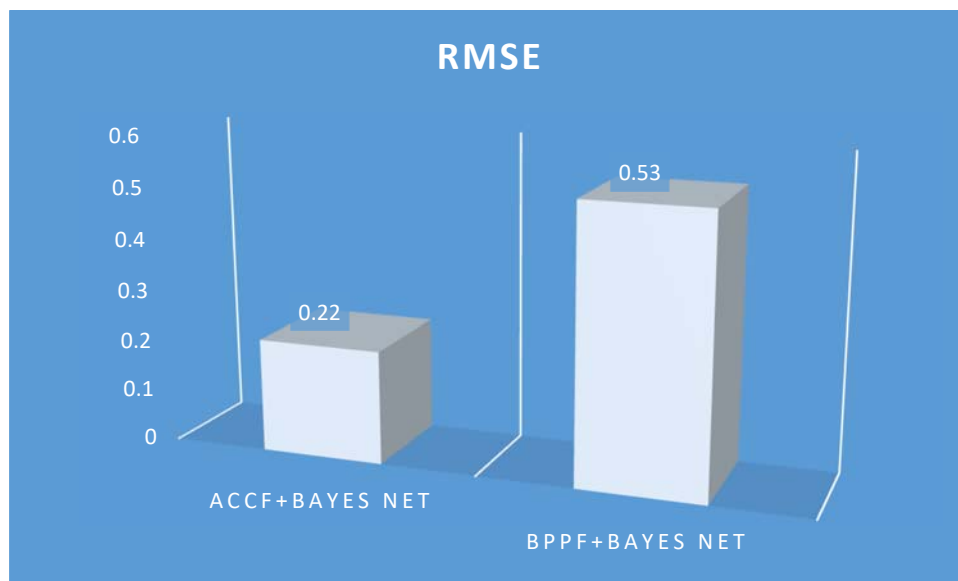

Figure 11: Representation of RMSE

The above diagram 48 shows that the Auto Color Correlogram Filter with Bayes Net gives root mean squared error 0.22, Binary Patter Pyramid Filter with Bayes Net gives RMSE 0.53. ACCF with BN performs well compared with BPPF with BN model.
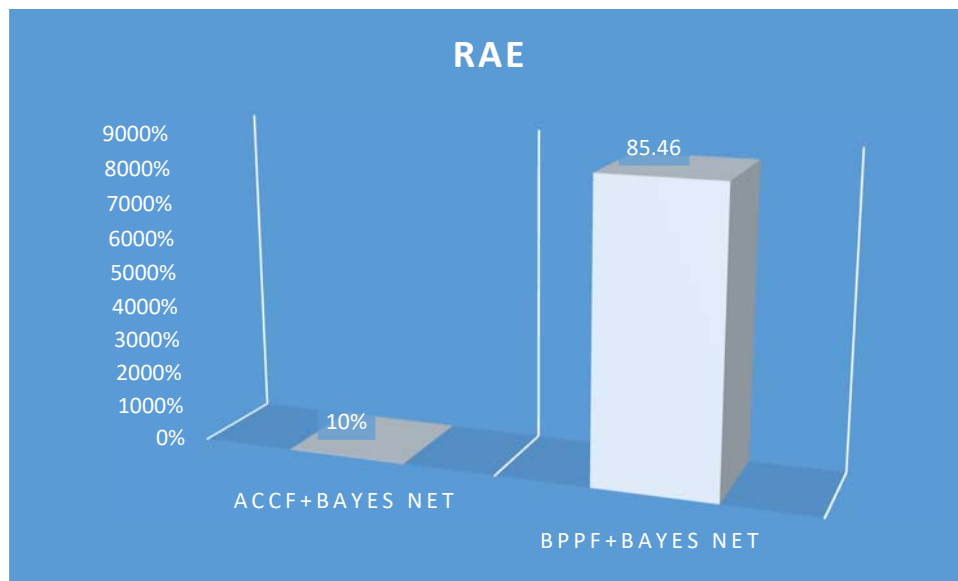
Figure 12: Representation of RAE

The above diagram 49 shows that the Auto Color Correlogram Filter with Bayes Net gives relative absolute error 10%, Binary Patter Pyramid Filter with Bayes Net gives RAE 85.46%. ACCF with BN performs well compared with BPPF with BN model.
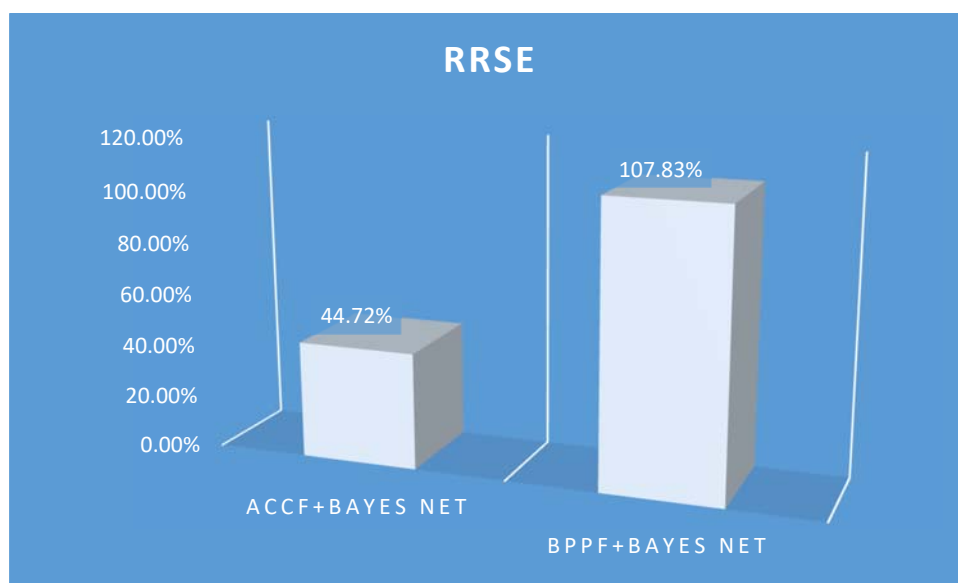

Figure 13: Representation of RRSE

The above diagram 50 shows that the Auto Color Correlogram Filter with Bayes Net gives root relative squared error 44.72%, Binary Patter Pyramid Filter with Bayes Net gives RRSE 107.83%. ACCF with BN performs well compared with BPPF with BN model.

Dr.G.Ayyappan et al. / Indian Journal of Computer Science and Engineering (IJCSE)
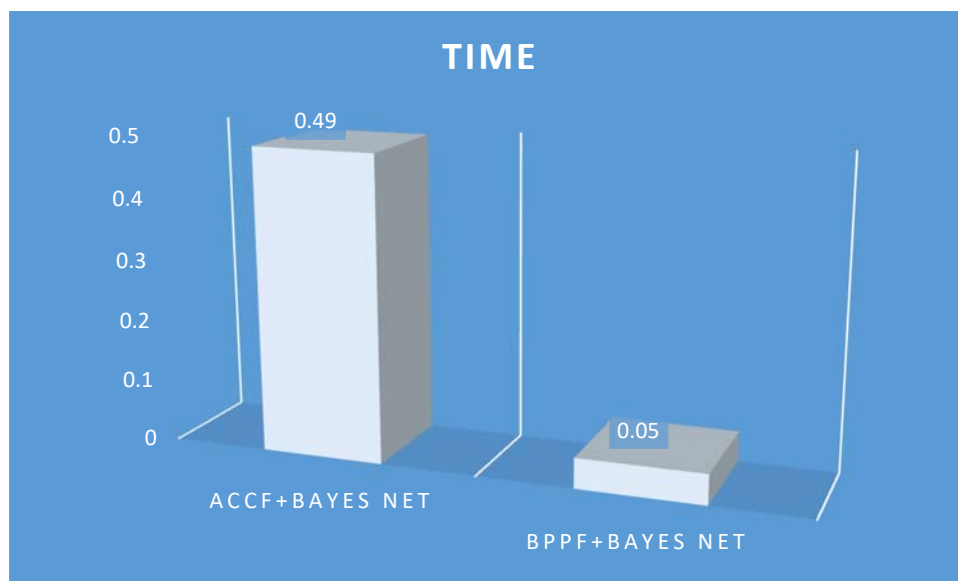

Figure 14: Representation of Time

The above diagram 51 shows that the Auto Color Correlogram Filter with Bayes Net model takes 0.49 seconds for creating its model, Binary Patter Pyramid Filter with Bayes Net takes 0.05 seconds for creating its model. APPF with BN performs well compared with ACCF with BN model.

IV Conclusion

This work concludes that the Auto Color Correlogram Filter with Bayes Net gives accuracy 95%, Binary Patter Pyramid Filter with Bayes Net gives accuracy 55%. The ACCF with BN performs well compared with BPPF with BN model.

Declarations

Conflicts of interest: The author's declare that they have no conflict of interest.

References

[1] Sumaira Ghazal, Waqar S. Qureshi, Umar S. Khan, Javaid Iqbal, Nasir Rashid, Mohsin I. Tiwana, Analysis of visual features and classifiers for Fruit classification problem, Computers and Electronics in Agriculture,Volume 187,2021,106267,ISSN 0168-1699, https://doi.org/10.1016/j.compag.2021.106267. (https://www.sciencedirect.com/science/article/pii/S0168169921002842)
[2] https://www.kaggle.com/datasets/karimabdulnabi/fruit-classification10-class
[3] https://www.cs.waikato.ac.nz/ml/weka/
[4] Abdel-Hakim, A.E., Farag A. A., 2006. CSIFT: A SIFT Descriptor with Color Invariant Characteristics. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 2006, pp. 1978-1983.
[5] Chamasemani, F. F., Singh, Y. P., 2011. Multi-class Support Vector Machine (SVM) Classifiers – An Application in Hypothyroid Detection and Classification. In: 2011 Sixth International Conference on Bio-Inspired Computing: Theories and Applications, Penang, 2011, pp. 351-356.
[6] Cunningham, P., Delany, S. J., 2007. K-Nearest neighbor classifiers. Technical Report UCD-CSI-2007-4, Dublin: Artificial Intelligence Group, March 27, 2007.
[7] Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, 2005, pp. 886–893.
[8] De Goma, J.C., Quilas, C.A.M., Valerio, M.A.B., Young, J.J.P., Sauli, Z., 2018. Fruit Recognition Using Surface and Geometric Information. J. Telecommunication Electronic Comput. Eng. (JTEC) 10 (1–15), 39–42.
[9] Fatima, S., Seshashayee, M., 2020. Healthy Fruits Image Label Categorization through Color Shape and Texture Features Based on Machine Learning Algorithm. Int. J. Innovative Technol. Exploring Eng. (IJITEE) 9 (3), 34–40.
[10] Graps, A., 1995. An Introduction to Wavelets. IEEE Comput. Sci. Eng. 2 (2), 50–61.
[11] Hossain, M.S., Al-Hammadi, M., Muhammad, G., Feb. 2019. Automatic Fruit Classification Using Deep Learning for Industrial Applications. IEEE Trans. Ind. Inf. 15 (2), 1027–1034.
[12] Katarzyna, R., Paweł, M., 2019. A Vision-Based Method Utilizing Deep Convolutional Neural Networks for Fruit Variety Classification in Uncertainty Conditions of Retail Sales. Appl. Sci. 9 (19).
[13] Jana, S., Parekh, R., 2017. Shape-based fruit recognition and classification. In: Proc. Int. Conf. Comput. Intell. Commun. Bus. Analytics, Mar. 2017, pp. 184-196.
[14] Kaviani, P., Dhotre, S., 2017. Short Survey on Naïve Bayes Algorithm. Int. J. Adv. Res. Comput. Sci. Manage. 4(11).
[15] Loesdau, M., Chabrier, S., Gabillon, A., 2014. Hue and Saturation in the RGB Color Space. In: Elmoataz A., Lezoray O., Nouboud F., Mammass D. (Eds.), Image and Signal Processing. ICISP 2014. Lecture Notes in Computer Science, vol. 8509, 2014, pp. 203-212.
[16] Lofstedt, ¨ T., Brynolfsson, P., Asklund, T., Nyholm, T., Garpebring, A., 2019. Gray-level invariant Haralick texture features. PLoS ONE 14 (2), 1–18.
[17] Macanh˜ a, P. A., Eler, D. M., Garcia, R. E., Junior, W. E., 2018. Handwritten feature descriptor methods applied to fruit classification. In: Latifi S. (eds) Information Technology - New Generations. Advances in Intelligent Systems and Computing, Cham: Springer, vol 558, 2018, pp. 699-705.

[18] Muresan, H., Oltean, M., Fruits 360 dataset on kaggle, [Online]. Available: https://www. kaggle.com/moltean/fruits.
[19] Muresan, H., Oltean, M., 2018. Fruit recognition from images using deep learning. Acta Univ Sapientiae, Informatica 10 (1), 26–42.
[20] Naranjo-Torres, J., Mora, M., Hern´ andez-García, R., Barrientos, R.J., Fredes, C., Valenzuela, A., May 2020. A Review of Convolutional Neural Network Applied to Fruit Image Processing. Appl. Sci. 10 (10), 3443.
[21] Nosseir, A., Ahmed, S.A., 2019. Automatic Classification for Fruits' Types and Identification of Rotten Ones using k-NN and SVM. Int. J. Online Biomed. Eng. 15 (3), 47–61.
[22] Patel, C. C., Chaudhari, V. K., 2019. Comparative Analysis of Fruit Categorization Using Different Classifiers. In: Advanced Engineering Optimization Through Intelligent Techniques, July 2019, pp. 153-164.
[23] Kumari, R. S. S., Gomathy, V., 2018. Fruit Classification using Statistical Features in SVM Classifier. In: 2018 4th International Conference on Electrical Energy Systems (ICEES), Chennai, 2018, pp. 526-529.
[24] Rojas-Aranda J. L., Nunez-Varela J. I., Cuevas-Tello J. C., Rangel-Ramirez G., 2020. Fruit Classification for Retail Stores Using Deep Learning. In; Mexican Conference on Pattern Recognition, June 2020, pp. 3-13.
[25] Sakib, S., Ashrafi, Z., Siddique, M. A., 2019. "Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers." arXiv, 2019, arXiv:1904.00783[cs.CV].
[26] Saranya, N., Srinivasan, K., Pravin Kumar, S. K., Rukkumani, V., Ramya, R., 2019. Fruit Classification Using Traditional Machine Learning and Deep Learning Approach. In; Computational Vision and Bio-Inspired Computing. ICCVBIC, pp. 79-89.
[27] Sazli, M.H., 2006. A brief review of feed-forward neural networks. Commun. Fac. Sci. Univ. Ank. Series A2–A3: Phys. Sci. Eng 50 (1), 11–17.
[28] Swain, P.H., Hauska, H., 1977. The decision tree classifier: Design and potential. IEEE Trans. Geosci. Electron. 15 (3), 142–147.
[29] Tharwat, A., Gaber, T., Ibrahim, A., Hassanien, A., 2017. Linear discriminant analysis: A detailed tutorial. AI Commun. 30, 169–190.
[30] Van de Sande, K.E.A., Gevers, T., Snoek, C.G.M., Feb. 2011. Empowering Visual Categorization With the GPU. IEEE Trans. Multimedia 13 (1), 60–70.
[31] Xiong, J., Liu, Z., Lin, R., Bu, R., He, Z., Yang, Z., Liang, C., March 2018. Green Grape Detection and Picking-Point Calculation in a Night-Time Natural Environment Using a Charge-Coupled Device (CCD) Vision Sensor with Artificial Illumination. Sensors 18 (4).

Authors Profile

Dr.G.Ayyappan, is working as an Associate Professor of Department of Information Technology at Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar, Chennai. His research interest includes Machine Learning, Image Processing, Social Networks, and Computational Intelligence. E-mail:ayyappangmca@gmail.com

Dr. R.Venkatesh Babu (V.B) is the Pro Vice- Chancellor of Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India. He has more than 23 years of experience in Teaching, Research, Academics and Administration. Having done his graduation in Mechanical Engineering, Post Graduation in Computer Aided Design, he did his Doctorate in the field of IC Engines. He is a member of the Academic Council, Board of Management of leading Universities in India. With more than 60 research articles in various indexed and peer reviewed national and international journals and conferences, 3 research scholars have been awarded Ph.D under his guidance and presently 4 are pursuing. His areas of research interest are AI, ML, Cyber Security, Quantum computing, FEA, CFD, MEMS, Smart Nano Materials, Supply Chain etc.As a professional Engineer, he is a Member of the American Society of Mechanical Engineers (ASME), Member of the Institution of Engineering and Technology (MIET), Life Member of the Indian Society for Technical Education (LMISTE) and Member of the Society of Automobile Engineers (SAE). Email Id:rvbaboo76@gmail.com

Dr.K.Alice is working as an Assistant Professor in SRMIST, Chennai, India. Her research interests are in Digital Image Processing, Machine Learning, Artificial Intelligence and information Security. E-mail: alicesureshme@gmail.com

Dr.A. Kumaravel is working as a Professor and Dean, School of Computing, Bharath Institute of Higher Education and Research, Chennai. His research interest includes Functional Paradigm, Soft Computing, Cloud Computing, Machine Learning, Computational Intelligence, and Knowledge Engineering. He is a life Member of ISTE and IET. He is a reviewer for IEEE transactions and Elsevier journals. E-mail: drkumaravel@gmail.com