

DOI: 10.38173/RST.2025.30.2.6:60-70

Title:	<i>STUDY THE DRYNESS OF HARVESTED TOMATO VIA IMAGE-MACHINE PROCESSING</i>
Authors:	Sharad Kumar OLI Saddam Husain DHOBI Kishori YADAV Suresh Prasad GUPTA

Section: Engineering
Issue: 2(30)/2025

Received: 24 April 2025	Revised: 8 July 2025
Accepted: 05 September 2025	Available Online: 11 September 2025

How to cite: Sharad Kumar OLI, Saddam Husain DHOBI, Kishori YADAV, Suresh Prasad GUPTA, “Study the Dryness of Harvested Tomato via Image-Machine Processing”, *Research and Science Today*, 2(30)/2025, pages 60-70, DOI: 10.38173/RST.2025.30.2.6:60-70

STUDY THE DRYNESS OF HARVESTED TOMATO VIA IMAGE-MACHINE PROCESSING

Sharad Kumar OLI¹
Saddam Husain DHOBI^{2*}
Kishori YADAV³
Suresh Prasad GUPTA^{4*}

ABSTRACT:

THIS STUDY INVESTIGATED THE DRYNESS OF TOMATOES THROUGH IMAGE PROCESSING AND MACHINE LEARNING TECHNIQUES. HISTOGRAM ANALYSIS REVEALED A CONSISTENT DECREASE IN INTENSITY OVER SEVEN DAYS POSTHARVEST, INDICATING A PROGRESSION TOWARD DRYNESS. IMPORTANTLY, HISTOGRAMS DEMONSTRATE VARYING DISTRIBUTIONS, WITH UNIFORM PATTERNS OBSERVED IN LATER DAYS, SUGGESTING A BALANCED MOISTURE CONTENT. MACHINE LEARNING ANALYSIS ASSESSES ACCURACY LEVELS ACROSS EPOCHS, WITH IMPORTED INDIAN TOMATOES EXHIBITING ACCURACIES RANGING FROM 0.48--0.52, WHEREAS LOCAL KATHMANDU TOMATOES RANGE FROM 0.47--0.51. NOTABLE TRENDS INCLUDE FLUCTUATIONS IN ACCURACY AND POTENTIAL OVERFITTING TENDENCIES, EMPHASIZING THE NEED FOR OPTIMIZATION AND STABILITY ENHANCEMENTS IN MACHINE LEARNING MODELS. A COMPARISON WITH PREVIOUS WORKS HIGHLIGHTS THE VARIABILITY IN MODEL PERFORMANCE, UNDERSCORING THE IMPORTANCE OF THOROUGH EXPERIMENTATION AND OPTIMIZATION. THE ABSTRACT ENCAPSULATES THE STUDY'S MAJOR FINDINGS, PROVIDING VALUABLE INSIGHTS INTO TOMATO DRYNESS ASSESSMENT THROUGH INTEGRATED IMAGE PROCESSING AND MACHINE LEARNING METHODOLOGIES.

KEY WORDS: IMAGE PROCESSING, MACHINE LEARNING TECHNIQUES, HISTOGRAM, INDIAN TOMATOES, KATHMANDU TOMATOES, DRYNESS

INTRODUCTION

Tomatoes can be dried via various techniques. However, the most common methods applied at the industrial scale are sun drying and convective hot air drying. Sun drying, the

¹Master in Physics, Department of Physics, Patan Multiple Campus, Tribhuvan University, Patan Dhoka Lalitpur Nepal, email: oli.sharadkumar@gmail.com

²Master in Physics, Department of Physics, Patan Multiple Campus, Tribhuvan University, Patan Dhoka Lalitpur Nepal, corresponding *email: saddam@ran.edu.np

³PhD in Physics, Department of Physics, Patan Multiple Campus, Tribhuvan University, Patan Dhoka Lalitpur Nepal, email: yadavkishori70@gmail.com

⁴PhD in Physics, Department of Physics, Patan Multiple Campus, Tribhuvan University, Patan Dhoka Lalitpur Nepal, corresponding *email: guptasirphys@gmail.com

traditional method prevalent in rural areas, involves natural dehydration under sunlight. However, compared with convective hot air drying, it has drawbacks. In the food industry, there is a preference for dryers ensuring better hygiene, improved nutritional value, and reduced drying time [1]. drying in hot air at relatively high temperatures reduces the presence of bacteria and fungi, increasing product preservation. Factors such as pretreatment, cultivar, and drying conditions influence dried tomato properties. Prolonged exposure to high temperatures and oxygen during hot-air drying leads to chemical reactions, affecting product quality in terms of composition, nutrition, structure, and flavor. Qualitative parameters such as color, shape, taste, and antioxidant content are also affected during the drying process [2].

Consumers prioritize color in dried tomatoes, which is highly affected by drying. Color change results from nonenzymatic browning and carotenoid degradation, which are caused mainly by browning, which impacts quality. With drying, moisture loss increases, exacerbated by temperature. High temperatures also degrade antioxidants such as ascorbic acid, which are naturally found in tomatoes [3]. Lycopene, a major carotenoid with health benefits, degrades during drying, affecting both color and nutrition. Heat and oxygen cause tissue destruction, further degrading lycopene. Isomerization also contributes to degradation, converting lycopene to a less biologically active form [4]. Intermittent or multistage drying offers economic and qualitative benefits compared with traditional methods. These processes involve altering the temperature, pressure, moisture, and air velocity during drying, leading to improved outcomes [5]. Temperature is a crucial factor, and a combination of high and low temperatures can shorten drying times while preserving product quality. Compared with continuous methods, multistage drying also saves energy. Optimizing the drying process reduces time and energy consumption and yields higher-quality products.

While many studies have examined the kinetics of moisture loss in dried tomatoes, only a few have explored the kinetics of qualitative parameters during drying. Additionally, research on modeling the changes in the physical parameters of tomatoes to optimize the drying process is lacking. Investigating physical parameters beyond moisture loss kinetics could enhance the development of multistage drying processes, offering quality improvements and significant reductions in drying times. Monitoring color, volume, and rehydration capacity during drying is crucial for industrial-scale implementation, as these parameters are easily determined and cost-effective [6]. Tomatoes play a vital role in agriculture and culinary fields, necessitating thorough quality assessment. This study aims to predict tomato quality by accurately assessing maturity levels and postharvest durations, which are specifically tailored to Sri Lankan market conditions, with a focus on Padma tomatoes. Using image processing techniques, maturity stages and postharvest dates are identified. Greenhouse-grown Padma tomatoes mimic market conditions for image capture, with convolutional neural networks aiding analysis. Model 1 achieves 99% accuracy in tomato classification, whereas Model 2 reaches 99% training accuracy and 98% validation accuracy. The integration of classification methods enhances tomato categorization, optimizing resource use [7].

Tomatoes, major vegetable and perennial fruit crops, are cultivated globally and are classified as fruits on the basis of their botanical characteristics. Despite their culinary classification as vegetables, tomatoes offer numerous health benefits because they are rich in nutrients, including vitamins, carotenoids, and phenolic compounds. They are widely used in processed products such as canned tomatoes, paste, and sauces. China, India, Turkey, and the United States lead in tomato production, with classifications based on factors such as maturity and color stages [8]. Tomatoes are highly valued horticultural commodities because of their economic importance and health benefits. They contain essential nutrients such as vitamins A, B, and C, as well as protein, carbohydrates, and minerals [9]. With widespread affordability,

tomatoes offer significant market potential accessible to all societal levels. Image processing plays a crucial role in analyzing tomato maturity on the basis of color images, contributing to efficient fruit quality assessment [10,11].

Despite extensive studies on various aspects of tomato drying processes, there remains a significant research gap in the area of incorporating image-machine processing techniques to assess the dryness of harvested mature tomatoes. While traditional methods rely on physical and chemical analyses, such as moisture content, color change, and nutrient degradation, there is a lack of research focused on the use of advanced image processing coupled with machine learning algorithms to evaluate the dryness of tomatoes. Specifically, there is a dearth of literature addressing the real-time assessment of tomato dryness via nondestructive methods, which could revolutionize quality control practices in the tomato processing industry. Additionally, while some studies have examined the maturation stages and postharvest durations of tomatoes via image processing, there is limited research on extending this approach to quantify dryness levels accurately. The significance of this study lies in its innovative approach to addressing a crucial research gap in tomato processing. By integrating advanced image processing and machine learning techniques, this study aims to revolutionize the assessment of tomato dryness, offering nondestructive, real-time monitoring capabilities. This novel methodology not only enhances accuracy and efficiency but also promises cost and resource savings while ensuring adaptability to diverse market conditions. Ultimately, this research has the potential to transform quality control practices in the tomato processing industry, leading to improved product quality, efficiency, and competitiveness in the market.

MATERIALS AND METHODS

For this study, two tomato samples were carefully chosen to ensure consistency and comparability throughout the experimental process. The selection criteria involved maintaining identical angles, positions, and environmental conditions for both samples over a period of seven days. One tomato was sourced as an imported sample from India, whereas the other was obtained locally from the market in Kathmandu. These tomatoes were carefully positioned to ensure uniformity in lighting and environmental factors across all observation days. By maintaining rigorous control over sample selection and preparation, this study aimed to minimize potential variables and focus specifically on the differences between imported and locally sourced tomatoes under identical conditions.

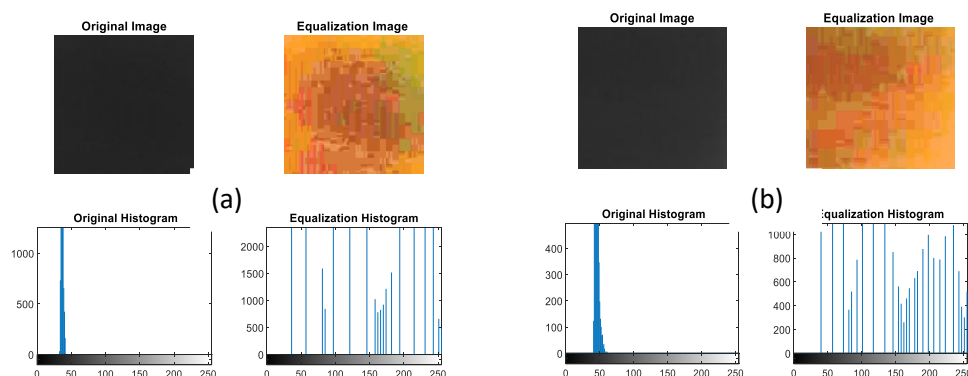
The research methodologies outlined involve distinct approaches tailored to their respective objectives within the fields of image processing [12] and machine learning [13]. For image processing, the methodology centers on conducting an experiment to evaluate the efficacy of histogram equalization in enhancing image contrast. This involves loading an image, applying histogram equalization, and visualizing the original and equalized images alongside their histograms. Additionally, a thresholding operation is performed to create a binary image, allowing for the observation of the impact of histogram equalization. This setup facilitates an investigation into the utility of histogram equalization for improving image contrast and aiding subsequent image processing tasks. On the other hand, the methodology for evaluating SVM classifiers' accuracy focuses on investigating the effect of varying training epochs on binary classification tasks using grayscale images. This methodology involves dataset selection, preprocessing, and the definition of an experimental setup with a specified number of images and epoch ranges. SVM classifiers are trained and evaluated across these epochs, with the dataset split into training and validation sets. Analysis involves plotting accuracy versus epochs to discern trends such as convergence or overfitting, leading to discussions on the optimal training epochs for SVM classifiers in binary classification tasks.

Both methodologies emphasize systematic experimentation, data analysis, and inference to address their respective research objectives in the image processing and machine learning domains.

In our investigation, we upheld the principles of repeatability and transparency by furnishing essential details about our dataset and execution environment. The study incorporates a training dataset alongside a testing dataset comprising 14 samples, maintaining consistent input–output structures for robust generalization assessment. All the experiments and simulations were conducted on a computer equipped with an Intel(R) Core(TM) i5-6200U CPU @ 2.30 GHz 2.40 GHz processor, utilizing MATLAB R2018a (version 9.4.0.183654). This choice ensured a stable environment conducive to data management, neural network modeling, and performance evaluation. Our meticulous documentation guarantees the reliability and authenticity of our study outcomes, facilitating reproducibility and further exploration of the proposed image-machine processing approach in analyzing the dryness of harvested mature tomatoes.

RESULTS AND DISCUSSION

This study aimed to analyze the dryness of harvested tomatoes on the basis of intensity and the number of pixels, employing histogram analysis through image processing techniques. Figures 1(a) to 1(g) depict the original histograms obtained over seven consecutive days following harvesting from the field. Our observations reveal a consistent decrease in intensity with an increase in the number of days postharvesting. Initially, as illustrated in Figure 1(a), a significant number of pixels with gray level intensities ranging between 35 and 45 were observed, exceeding 1000. This indicated a high level of moisture content in the freshly harvested tomatoes. However, as the days progressed, the intensity distribution shifted toward higher gray levels, indicating a decrease in moisture content and the onset of dryness. For example, in Figure 1(g), the number of pixels with gray-level intensities between 60 and 120 reached 225, indicating a reduction in moisture content compared with earlier days. Furthermore, the equalization histograms, represented in Figures 1(a) to 1(g), provided insights into the distribution of gray levels. Notably, a uniform distribution was observed in the histograms corresponding to the last two days of the experiment, suggesting a balanced intensity distribution as the tomatoes approached a drier state. Conversely, uneven distributions were observed in the histograms of the earlier days, indicating varying moisture contents.



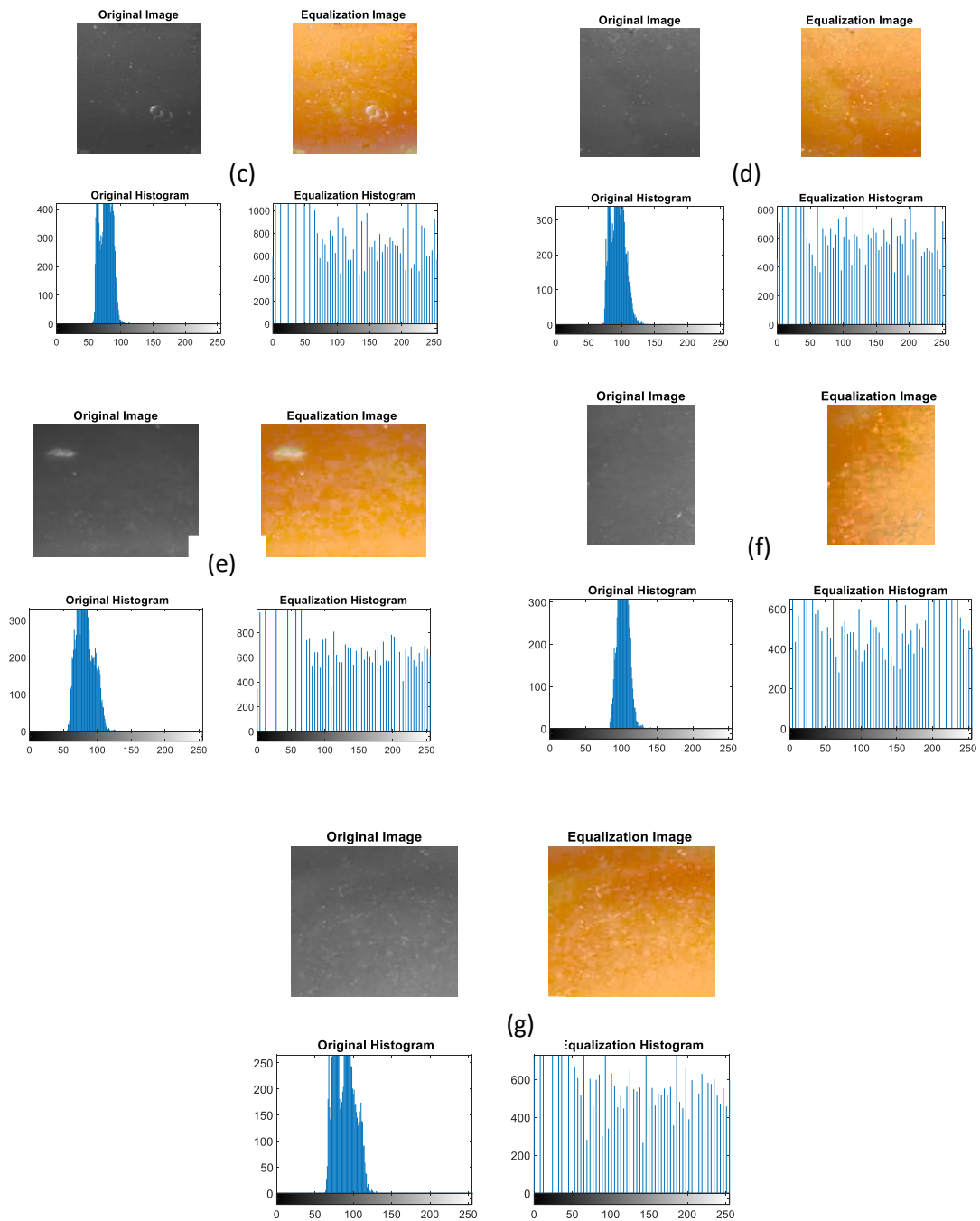
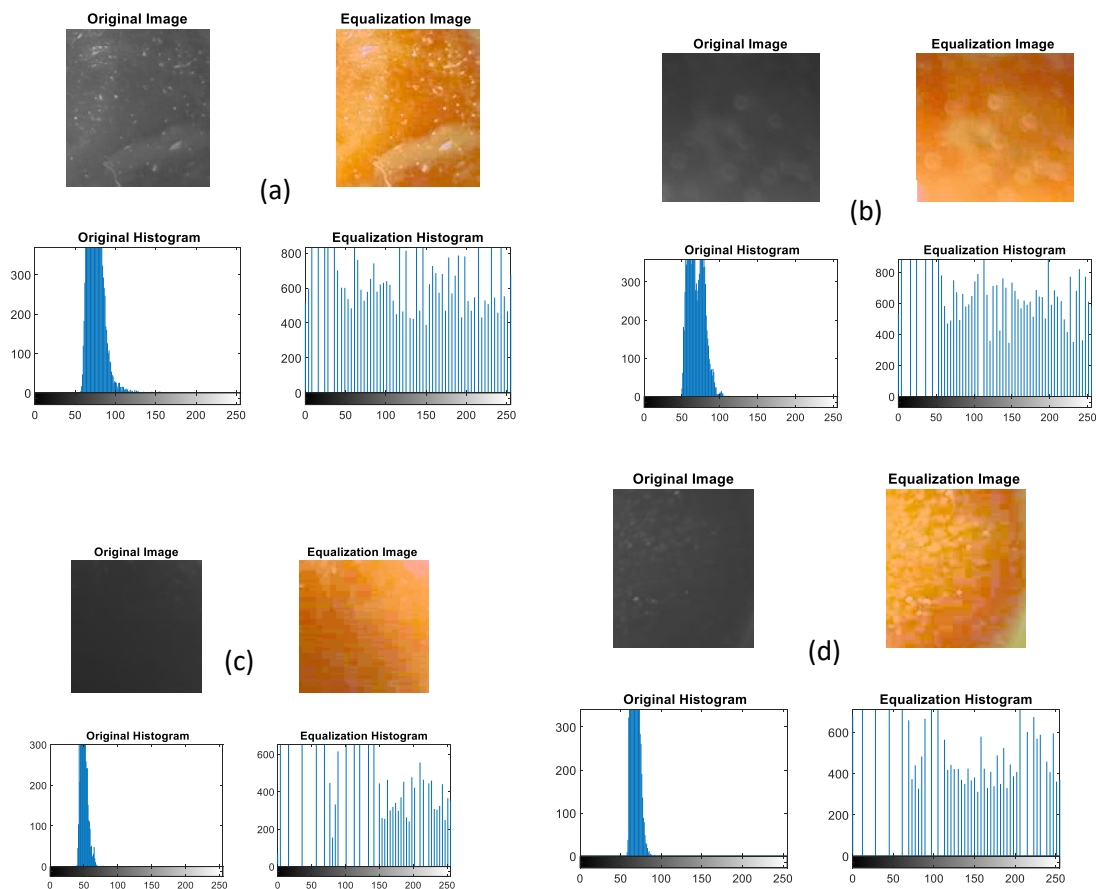


Figure 1: Intensity vs. Pixel of tomato imported from India

Specifically, the histograms of the first and second days (Figures 1(a) and 1(b)) exhibited similar patterns, with a predominant concentration of pixels at a gray level intensity of 2000 and a nonuniform distribution, respectively. Similarly, the histograms of the third and fourth days (Figures 1(c) and 1(d)) displayed analogous distributions, albeit with a higher intensity in the middle for the former, whereas the latter exhibited a more uniform distribution at a gray level intensity of 600. Similarly, the intensity distributions of the fifth and sixth days (Figures 1(e) and 1(f)) were characterized by uniform distributions at gray-level intensities of 600 and 400, respectively. Finally, the histogram of the seventh day (Figure 1(g)) depicted a

uniformly distributed intensity at a gray level of 600, indicating a more generalized dryness state. These findings suggest a progressive decrease in moisture content and an increase in dryness as the days postharvesting advance. The observed patterns in the histograms provide valuable insights into the dynamic changes in the intensity and distribution of gray levels, contributing to a deeper understanding of the dryness process in harvested tomatoes.

The original histograms obtained during image processing of tomatoes harvested from the field over seven days are depicted in Figures 2(a) to 2(g). The observations revealed a consistent decrease in intensity as the days postharvest progressed, indicating a reduction in moisture content and an increase in dryness. Initially, as shown in Figure 2(a), a substantial number of pixels with gray level intensities ranging between 50 and 100 were observed, exceeding 350. This high intensity suggests a relatively high moisture content in the freshly harvested tomatoes. However, as the days advance, the intensity distribution shifts toward lower gray levels, indicating a decrease in moisture content and progression toward a drier state. For example, in Figure 2(g), the number of pixels with gray-level intensities between 50 and 80 reaches 250, indicating a reduction in moisture content compared with earlier days. Moreover, the equalization histograms, represented in Figures 2(a) to 2(g), provide insights into the distribution of gray levels. Notably, a uniform distribution is observed in the histograms corresponding to the last two days of the experiment, suggesting a balanced intensity distribution as the tomatoes approach a drier state. Conversely, uneven distributions are observed in the histograms of the earlier days, indicating varying moisture contents.



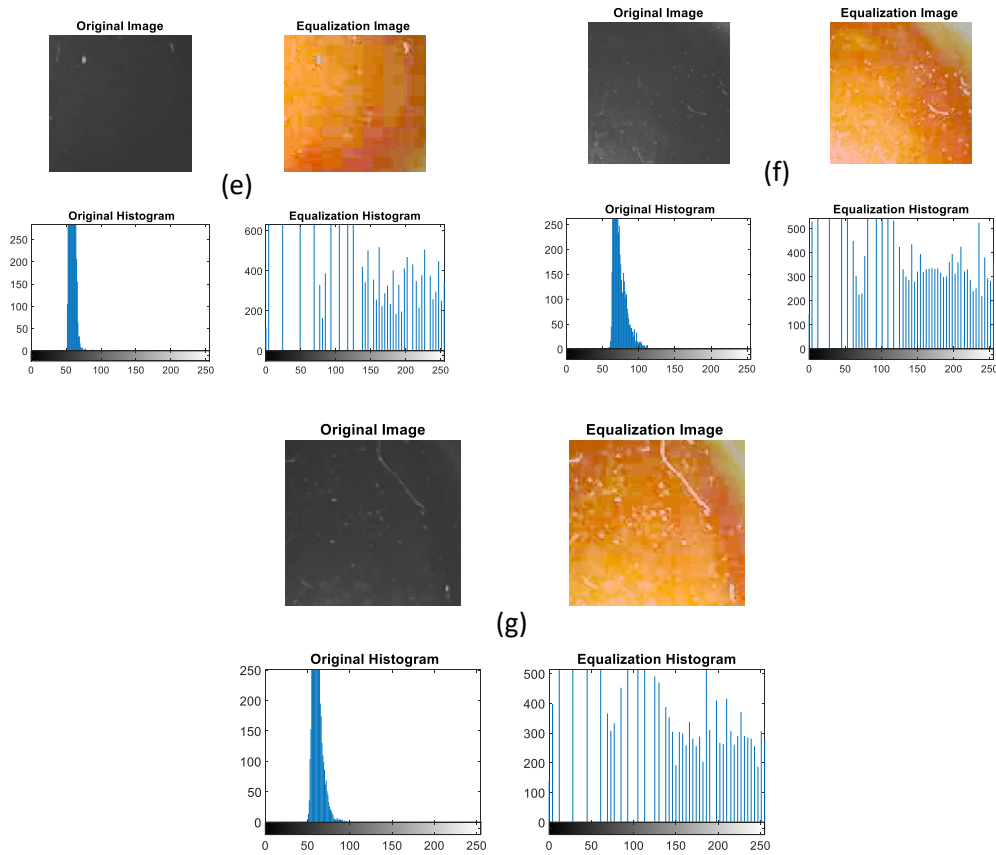


Figure 2: Intensity vs. pixel of tomato taken from the local market of Nepal (Kathmandu)

Specifically, the histograms of the first and second days (Figures 2(a) and 2(b)) exhibit similar patterns, with a predominant concentration of pixels at gray level intensity 600 but not uniformly distributed. Similarly, the histograms of the third and fourth days (Figures 2(c) and 2(d)) display analogous distributions, albeit with a higher intensity in the middle for the former, whereas the latter exhibits a more uniform distribution at a gray level intensity of 400 but not uniformly. Likewise, the intensity distributions of the fifth and sixth days (Figures 2(e) and 2(f)) are characterized by uniform distributions at gray-level intensities of 200 and 300, respectively. Finally, the histogram of the seventh day (Figure 2(g)) depicts a uniformly distributed intensity at a gray level of 250, indicating a more generalized dryness state. These findings suggest a progressive decrease in moisture content and an increase in dryness as the days postharvesting advance. The observed patterns in the histograms provide valuable insights into the dynamic changes in the intensity and distribution of gray levels, contributing to a deeper understanding of the dryness process in harvested tomatoes from the local market of Nepal (Kathmandu).

The application of image-machine learning techniques for assessing the dryness of tomatoes yielded varying levels of accuracy across different epochs. The accuracy levels for exported Indian tomatoes ranged from 0.48--0.52, indicating moderate performance of the machine learning model. Figure 3 illustrates the trends of the training and validation accuracies with respect to the number of epochs, and the natural accuracy validates the results (Abekoon et al., 2024). Upon analyzing the observations, distinct patterns emerge in the validation and training accuracies across different days of observation. For the first day, the validation

accuracy consistently surpasses the training accuracy for up to 40 epochs. This trend suggests that the model generalizes well to unseen data, indicating robust performance. Conversely, on the second day, the validation accuracy consistently outperforms the training accuracy across all epochs, indicating that the model effectively learns from the training data and generalizes to the validation set. However, on the third day, the model's performance shifts, with the training accuracy consistently surpassing the validation accuracy across all epochs. This discrepancy suggests potential overfitting, where the model may capture noise or specific patterns in the training data that do not generalize well to unseen data.

The observations for the fourth and fifth days reveal fluctuations in both the training and validation accuracies, indicating instability in the model's performance. This inconsistency may arise from various factors, such as insufficient training data, model complexity, or inadequate regularization techniques. Similarly, on the sixth day, although the overall trend shows reasonable performance, fluctuations are observed, suggesting the need for further optimization or fine-tuning of the model parameters. Finally, on the seventh day, the training accuracy consistently exceeds the validation accuracy across all epochs. This discrepancy may indicate a potential bias toward the training data, suggesting the need for additional data augmentation or regularization techniques to improve the generalization performance.

Comparing these results with those of previous works, it is evident that the performance of the machine learning model in assessing tomato dryness varies depending on the dataset characteristics, model architecture, and optimization techniques employed. While some studies may report higher accuracy levels, others may encounter challenges similar to those observed in this study, such as overfitting, fluctuations in accuracy, or bias toward the training data. These findings highlight the importance of thorough experimentation, model evaluation, and optimization to increase the accuracy and reliability of image-machine learning systems for assessing the dryness of harvested tomatoes. Additionally, future research could focus on exploring advanced model architectures, incorporating domain-specific features, and employing regularization techniques to improve the generalization performance and robustness.

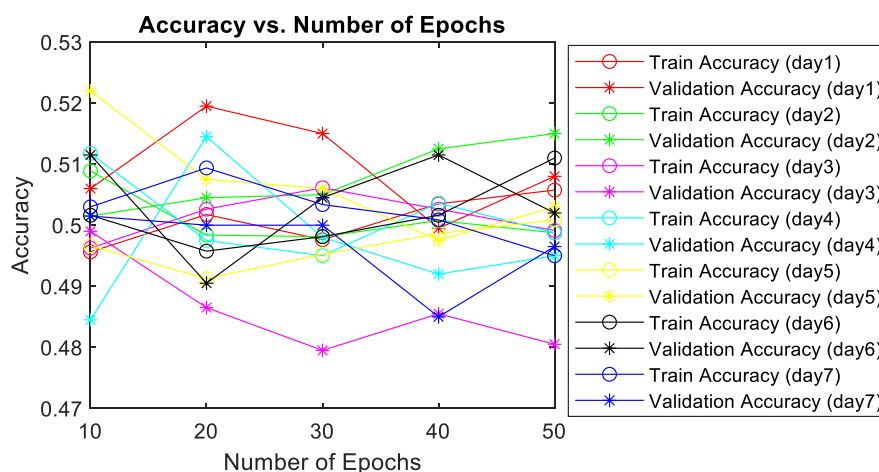


Figure 3: Intensity vs. Pixel of tomato imported from India

The evaluation of local tomatoes from Kathmandu, Nepal, via image-machine learning techniques revealed accuracy levels ranging from 0.47--0.51 across different epochs. Figure 4 depicts the trends of the training and validation accuracies with respect to the number of epochs. Upon scrutinizing the observations, notable patterns emerge in the validation and training

accuracies across the seven days of observation. On the first day, the training accuracy consistently exceeds the validation accuracy up to 40 epochs before fluctuating at certain values. This trend suggests that the model may initially overfit the training data but may exhibit instability as the number of epochs increases.

Conversely, on the second and third days, the validation accuracy consistently outperforms the training accuracy across all the epochs. This indicates effective generalization of the model to unseen data, suggesting robust learning capabilities and a reliable assessment of tomato dryness. However, on the fourth and fifth days, both the training and validation accuracies exhibit high fluctuations across all the epochs. This instability may stem from various factors, such as dataset variability, model complexity, or insufficient regularization techniques. These fluctuations underscore the need for further optimization and stability enhancements in the machine learning model. On the sixth day, while the overall trend suggests reasonably consistent performance, fluctuations are observed, particularly at certain epochs. This variability may warrant additional investigation into model architecture, hyperparameter tuning, or dataset augmentation to improve stability and generalization performance. Finally, on the seventh day, the training accuracy consistently surpasses the validation accuracy above 20 epochs. This discrepancy suggests potential overfitting to the training data, indicating the need for regularization techniques or data augmentation strategies to enhance generalizability.

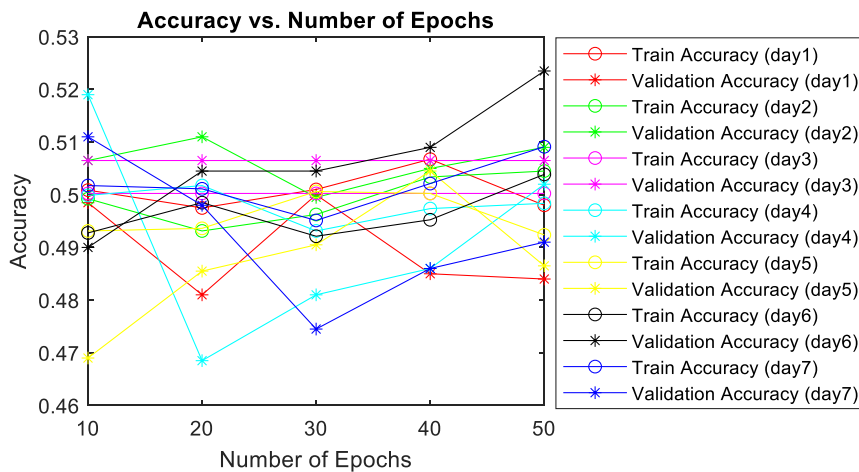


Figure 4: Intensity vs. Pixel of tomatoes taken from the local market of Nepal (Kathmandu)

Comparing these findings with those of previous works, it is evident that the performance of the machine learning model in assessing tomato dryness varies depending on factors such as dataset characteristics, model architecture, and optimization techniques employed. While some studies may report higher accuracy levels and stable performance, others may encounter challenges such as fluctuations in accuracy or overfitting tendencies. These observations underscore the importance of comprehensive experimentation, model evaluation, and optimization to increase the accuracy and reliability of image-machine learning systems for assessing tomato dryness. Additionally, future research could focus on refining model architectures, incorporating domain-specific features, and employing regularization techniques to improve stability and generalization performance across diverse datasets.

CONCLUSION

This study comprehensively examines the dryness of harvested tomatoes via integrated image processing and machine learning techniques. Through histogram analysis, a progressive

decrease in intensity over seven days postharvest is observed, indicating a transition toward dryness. Machine learning analysis reveals varying accuracy levels across epochs, with imported Indian tomatoes exhibiting accuracies between 0.48 and 0.52 and local Kathmandu tomatoes ranging from 0.47 to 0.51. Notable trends include fluctuations in accuracy and potential overfitting tendencies, emphasizing the need for optimization and stability enhancements in machine learning models. Overall, the study underscores the effectiveness of combined image processing and machine learning methods in assessing tomato dryness, offering valuable insights for the agricultural and food processing industries. These findings contribute to the advancement of techniques for quality control and preservation of harvested produce, paving the way for future research in optimizing and refining such methodologies for broader applications in agricultural science and food technology.

REFERENCES

- [1]. A. Dauda, O. Abiodun, T. Salami, and O. Akintayo, "Chemical and microbiological evaluation of dried tomato slices for Nigerian system," *Glob. J. Nutr. Food Sci.*, 2019, doi: 10.33552/gjnfs.2019.01.000521.
- [2]. B. Zanoni, C. Peri, R. Nani, and V. Lavelli, "Oxidative heat damage of tomato halves as affected by drying," *Food Res. Int.*, vol. 31, no. 5, pp. 395–401, 1999, doi: 10.1016/S0963-9969(98)00102-1.
- [3]. S. Cernișev, "Effects of conventional and multistage drying processing on non-enzymatic browning in tomato," *J. Food Eng.*, vol. 96, no. 1, pp. 114–118, 2010, doi: 10.1016/j.jfoodeng.2009.07.002.
- [4]. J. Shi, M. le Maguer, Y. Kakuda, A. Liptay, and F. Nickamp, "Lycopene degradation and isomerization in tomato dehydration," *Food Res. Int.*, vol. 32, pp. 15–21, 1999, doi: 10.1016/S0963-9969(99)00059-9.
- [5]. S. J. Kowalski and A. Pawłowski, "Energy consumption and quality aspect by intermittent drying," *Chem. Eng. Process.*, vol. 50, no. 4, pp. 384–390, 2011, doi: 10.1016/j.cep.2011.02.012.
- [6]. I. Doymaz and Ö. Özdemir, "Effect of air temperature, slice thickness and pretreatment on drying and rehydration of tomato," *Int. J. Food Sci. Technol.*, vol. 49, no. 2, pp. 558–564, 2014, doi: 10.1111/ijfs.12337.
- [7]. T. Abekoon, H. Sajindra, J. A. D. C. A. Jayakody, E. R. J. Samarakoon, and U. Rathnayake, "Image processing techniques to identify tomato quality under market conditions," *Smart Agric. Technol.*, vol. 7, 100433, 2024.
- [8]. Y. Li, H. Wang, Y. Zhang, and C. Martin, "Can the world's favorite fruit, tomato, provide an effective biosynthetic chassis for high-value metabolites?," *Plant Cell Rep.*, vol. 37, no. 10, pp. 1443–1450, 2018, doi: 10.1007/s00299-018-2283-8.
- [9]. Tugiyono, *Tanaman Tomat*. Jakarta: Agromedia Pustaka, 2005.
- [10]. A. Basuki, J. F. Palandi, and Fatchurrochman, *Pengolahan Citra Digital Menggunakan Visual Basic*. Yogyakarta: Graha Ilmu, 2005.
- [11]. S. Apriani, *Penerapan Image Processing untuk Mengetahui Tingkat Kematangan Buah Belimbing Manis (Averrhoa Carambola L.)*. Palembang: Universitas Sriwijaya, 2017.
- [12]. A. Priyono, A. Rejo, and E. A. Kuncoro, *Aplikasi Teknologi Pengolahan Citra pada Perubahan Warna Buah Jeruk Selama Proses Degreening*. Skripsi, Universitas Sriwijaya, Indralaya, 2008.
- [13]. C. H. Putri, A. Rejo, and E. A. Kuncoro, *Penerapan Teknik Image Processing untuk Menentukan Tingkat Kematangan Buah Tomat*. Skripsi, Universitas Sriwijaya, Palembang, 2003.