International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 2, Mar-Apr 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Detection of Chemical and Natural Ripened Banana Using Image-Based Machine Learning

BG Varsha*, Abhinaya**, MM Aishwarya, B Sushmita

*(Department of Computer Science and Engineering — Artificial Intelligence, Ballari Institute of Technology and Management, Ballari, Karnataka, India
Email: varsha97415@gmail.com)

** (Department of Computer Science and Engineering — Artificial Intelligence, Ballari Institute of Technology and Management, Ballari, Karnataka, India
Email: abhinayagouda28@gmail.com)

--************************----------------------------------

Abstract:

Fruit safety at the point of consumer purchase has turned into a serious challenge across global food supply networks. Bananas (Musa acuminata) are among the most widely consumed fruits in South and Southeast Asia, yet their quality is frequently compromised by unauthorised use of calcium carbide as an artificial ripening agent. When this compound meets atmospheric moisture it gives off acetylene gas, triggering rapid peel yellowing while depositing arsenic and phosphorus hydride residues that are toxic to human health. Buyers at retail markets have no straightforward way to tell treated fruit from untreated fruit with the naked eye, and laboratory-based confirmation methods are too slow, costly, and destructive for routine point-of-sale use. This paper describes a Convolutional Neural Network (CNN) system that classifies banana photographs as either naturally ripened or chemically ripened. A labelled dataset of 4,000 banana images was built from retail markets, wholesale centres, and controlled laboratory ripening experiments. Images were standardised to 128×128 pixels, pixel values were scaled to [0,1], and a real-time augmentation routine applied random rotations, flips, zoom, brightness shifts, and spatial translations during training. A six-layer sequential CNN trained with the Adam optimiser under binary cross-entropy loss reached 93% overall classification accuracy on a separate test set, with balanced precision and recall near 0.93 for both classes. The findings confirm that deep learning applied to standard RGB photographs offers a practical, non-destructive, and inexpensive screening option for banana safety assessment, with direct scope for integration into mobile consumer tools and automated supply-chain inspection platforms..

Keywords — **Banana ripeness detection, convolutional neural network, image classification, chemical ripening, calcium carbide, deep learning, food safety, TensorFlow, OpenCV, computer vision**

--************************----------------------------------

I. INTRODUCTION

Fresh produce quality control has always depended heavily on the skill and experience of individual inspectors. In large, fast-moving supply chains this creates obvious weaknesses: throughput is limited, fatigue introduces inconsistency, and genuinely subtle defects are easy to miss. Automated image analysis offers a way to address each of these weaknesses simultaneously, and the falling cost of camera hardware together with the maturity of open-source deep learning libraries has made practical deployment increasingly feasible.

Among the fruits that flow through South Asian markets in high volume, bananas occupy an especially prominent position. World annual production has surpassed 120 million metric tonnes, with India, the Philippines, and Ecuador among the leading producers. In India in particular, bananas are

 ©IJSRED: All Rights are Reserved

consumed at every socioeconomic level and represent a significant income source for smallholder farmers. Their composition — potassium, B-group vitamins, dietary fibre, and rapidly available sugars — makes them a nutritionally important staple food rather than merely a discretionary purchase.

Left on the tree or in a cool store, a banana completes its natural ripening in roughly seven to ten days after harvest. The hormone ethylene, produced within the fruit tissue itself, gradually breaks down chlorophyll so the peel transitions from green to yellow, while amylase enzymes simultaneously convert stored starch into the free sugars that give ripe bananas their characteristic sweetness. At full natural ripeness a scatter of brown flecks typically appears on the peel, the outward sign of continued enzymatic activity beneath the surface.

Market economics create persistent pressure to compress that ten-day window. A trader who cannot carry the cost of waiting, or who needs to synchronise delivery with peak retail demand, may expose unripe bananas to calcium carbide CaC_2). In the presence of moisture this compound releases acetylene (C_2H_2), a gas that mimics ethylene and drives visible yellowing within 24 to 48 hours. The problem is that commercial-grade calcium carbide is rarely pure. It carries arsenic trioxide and phosphorus hydride as manufacturing by-products, and both of these substances transfer to the fruit surface during treatment. Regular ingestion of such residues has been linked to headache, nausea, gastrointestinal inflammation, liver stress, and — with longer-term exposure — neurological damage and elevated cancer risk. Calcium carbide ripening is explicitly banned under the Food Safety and Standards Act in India and under comparable legislation in the European Union and the United States, yet enforcement at the market level remains inconsistent.

The difficulty for ordinary buyers is that a chemically ripened banana and a naturally ripened one can look almost the same under typical shop lighting. The chemically treated fruit tends to be uniformly bright yellow with no brown spots, and sometimes retains patches of green near the tip, but these differences are not always obvious. Standard confirmatory tests — silver nitrate spot testing, acidified potassium permanganate, gas chromatography — are accurate but share a common set of drawbacks: they need laboratory space and trained operators, they take hours per sample, and they destroy the fruit being tested, making them useless for screening produce that is still for sale.

A CNN-based image classifier sidesteps all those constraints. The network processes a single photograph taken with a smartphone and returns a binary label: naturally ripened or chemically ripened. CNNs learn to extract relevant spatial features — colour uniformity, peel texture, spot patterns, surface reflectance — directly from raw pixel data, with no hand-crafted feature engineering required. Once trained, the model runs in tens of milliseconds on commodity hardware, at zero per-sample cost, and without touching the fruit.

The contributions of this study are: construction of a 4,000-image labelled dataset spanning both ripening categories collected across real market and controlled laboratory settings; design and training of a six-layer sequential CNN optimised for the binary classification task; rigorous evaluation on a held-out test partition using accuracy, precision, recall, and F1-score; and a frank discussion of practical limitations alongside concrete directions for future extension.

II. Literature Review

A. Early Image-Processing Approaches to Fruit Quality

Before deep learning became the dominant paradigm, researchers characterised fruit quality from digital images through manually constructed feature pipelines. Colour histograms in RGB and L*a*b* spaces, co-occurrence matrix texture statistics, and basic shape descriptors were extracted and fed to support vector machines, k-nearest-neighbour classifiers, or decision trees. These pipelines worked reasonably well inside the controlled imaging environments used for training but generalised poorly when lighting conditions, camera angles, or fruit varieties shifted outside the narrow range covered by the training data. Reported accuracies hovered in the 70 to 82 percent range, with variance that made routine deployment unreliable [4].

B. Convolutional Networks in Agricultural Image Analysis

The publication of Mohanty et al.'s [1] PlantVillage study in 2016 established a new benchmark for the field. Training a CNN on over 54,000 leaf images, they achieved better than 99 percent accuracy in identifying 26 diseases across 14 crop species, demonstrating conclusively that learned feature hierarchies could outpace handcrafted ones on agricultural imagery. Sa et al. [2] extended this result to uncontrolled outdoor environments, detecting and localising multiple fruit types at different growth stages despite variable illumination and partial occlusion. The foundational innovations that made such results possible — ReLU activations, dropout regularisation, augmentation during training — were introduced to the wider community by Krizhevsky, Sutskever, and Hinton [3] with the AlexNet architecture and have since become standard components in virtually every CNN built for image classification.

C. Ripeness Grading of Bananas Specifically

Research targeting banana ripeness in particular began with colour-chart comparisons, mapping peel hue against the seven-stage Banana Ripeness Colour Chart. These colour-only methods were easy to implement but broke down under artificial lighting. Incorporating L*a*b* colour statistics alongside GLCM-derived texture features and using ensemble classifiers improved accuracy to roughly 87 to 89 percent on laboratory datasets. Bhargava and Bhargava [6] demonstrated that even shallow CNNs trained end-to-end on banana and other fruit images outperformed these feature-engineering pipelines, reaching 90 to 92 percent for freshness grading across multiple categories. LeCun, Bengio, and Hinton [5] provided the broader theoretical grounding for why hierarchical feature learning works so well on visual data.

D. Detecting Chemical Treatment Specifically

Studies aimed specifically at identifying calcium carbide treatment, rather than general ripeness stage, are sparse. Sirimanna et al. [7] tackled this problem for mangoes, using SVM classification of peel colour statistics to reach about 85 percent accuracy and identifying colour uniformity and the absence of natural ripening spots as the two most informative visual cues. Islam et al. [4] built a machine learning pipeline for chemically ripened fruit across several species, achieving 83 to 88 percent with traditional feature engineering and explicitly noting that a shift to deep learning was the logical next step for improved performance. Hyperspectral and near-infrared techniques offer higher sensitivity to internal chemical markers but require instrumentation that is too expensive and physically bulky for field screening [10]

E. Transfer Learning for Small Food-Safety Datasets

Transfer learning — initialising a network with weights from a large general-purpose dataset and fine-tuning on a smaller domain-specific one — has proven especially valuable when labelled training data is limited. Architectures such as VGG-16 [11], ResNet-50 [12], MobileNetV2, and EfficientNet have each been successfully adapted to food quality tasks, consistently bettering from-scratch models by four to ten percentage points in accuracy while converging faster. The advantage is largest precisely when the domain dataset is in the hundreds to low thousands of images — the scale relevant to the present study.

E. Gap This Study Fills

The existing body of work has not produced a purpose-built CNN system that tackles the specific binary problem of naturally versus chemically ripened bananas using only standard RGB images, evaluated against a dedicated labelled dataset collected under real market conditions. Equally, few prior systems were designed with practical consumer deployment in mind, accounting for constraints such as smartphone hardware, fast inference, and usability by people without technical backgrounds. This study addresses both points directly.

III. METHODOLOGY

Four stages make up the experimental pipeline: dataset construction, image preprocessing, CNN architecture design, and model training with evaluation. The system was built to be reproducible and to support straightforward extension to additional fruit categories in future work.

A. Dataset Construction

 ©IJSRED: All Rights are Reserved

No publicly available benchmark captures the naturally-versus-chemically-ripened distinction for bananas under typical Indian market conditions, so a purpose-built labelled dataset was assembled. Images came from four source types: open-air retail fruit markets in Ballari and surrounding areas; wholesale agricultural depots; controlled laboratory settings where bananas were ripened either under ambient ethylene or under standardised calcium carbide exposure; and publicly accessible online fruit image repositories. Drawing from all four environments was deliberate, ensuring the dataset would cover the range of lighting, camera distances, viewing angles, and peel presentations a deployed system would face in practice.

Naturally ripened examples show bananas that matured under ambient temperature and humidity over seven to ten days. Their images display a gradual green-to-yellow gradient, scattered brown flecks, and a slightly uneven matte peel texture. Chemically ripened examples depict fruit exposed to calcium carbide gas for 24 to 48 hours: sudden uniform yellowing, absent or minimal brown speckle, a smoother sometimes slightly waxy surface, and occasional residual green near the tip or crown. Two domain experts — an agricultural scientist and a food safety officer — labelled each image independently, with disagreements resolved by discussion.

The finished dataset holds 2,000 images per class, giving a balanced total of 4,000 annotated images. Stratified random sampling divided the data into training (70%, n=2,800), validation (15%, n=600), and test (15%, n=600) subsets while preserving the 50:50 class ratio across all three partitions.

B. Image Preprocessing

Every image passed through a consistent pipeline before model training or inference. Raw resolutions in the dataset ranged from 480×640 to 3024×4032 pixels depending on the capture device; all images were resized to 128×128 pixels via bilinear interpolation. This target resolution preserves enough peel detail to support accurate classification while keeping memory and computation costs manageable.

Pixel intensities originally in the integer range [0, 255] were divided by 255 to produce floating-point values in [0.0, 1.0]. This rescaling puts all input features on a comparable numerical scale, which stabilises gradient dynamics and speeds convergence during stochastic optimisation.

Real-time data augmentation was applied only to the training partition during each epoch, using the Keras ImageDataGenerator class. Transformations included: horizontal and vertical flipping at 50 percent probability each; rotation within ±30 degrees; zoom within the range [0.8, 1.2]; brightness adjustment within [0.7, 1.3]; and width and height shifts within 10 percent of image dimensions. Applying these transformations afresh each epoch multiplies effective dataset diversity without increasing storage requirements, training the network to stay invariant to the incidental variation in framing, lighting, and angle that characterises real-world capture.

C. CNN Architecture

The network follows a sequential structure: three convolutional extraction blocks each pairing a Conv2D layer with a max-pooling layer, followed by a fully connected classification head. Filter counts increase across blocks — 32, 64, 128 — following the established principle that early layers should capture generic low-level features while later layers encode increasingly discriminative task-specific representations. All convolutions use 3×3 kernels with same-mode padding and ReLU activation; each pooling operation uses a 2×2 window with stride 2. After the third pooling step the 16×16×128 feature map is flattened to 32,768 values, passed through a 256-unit Dense layer with ReLU, regularised by a Dropout layer (rate 0.5), and finally projected to a single sigmoid output neuron whose value is thresholded at 0.5 for class assignment. The complete layer specification is given in Table I.

TABLE I

PROPOSED CNN ARCHITECTURE — LAYER-BY-LAYER SUMMARY

Layer	Configuration	Output Shape
Input	128×128×3 RGB image	128×128×3

 ©IJSRED: All Rights are Reserved

Layer	Configuration	Output Shape
Conv2D + ReLU	32 filters, 3×3, same padding	128×128×32
MaxPooling2D	2×2 pool, stride 2	64×64×32
Conv2D + ReLU	64 filters, 3×3, same padding	64×64×64
MaxPooling2D	2×2 pool, stride 2	32×32×64
Conv2D + ReLU	128 filters, 3×3, same padding	32×32×128
MaxPooling2D	2×2 pool, stride 2	16×16×128
Flatten	3D feature map → 1D vector	32,768
Dense + ReLU	256 fully connected units	256
Dropout	Rate = 0.5	256
Dense + Sigmoid	1 unit, binary output	1

D. Training Configuration

The model was compiled with the Adam optimiser at an initial learning rate of 0.001 and binary cross-entropy loss. Adam was chosen for its adaptive per-parameter moment estimates, which provide consistent convergence across a wide range of CNN training scenarios without extensive manual tuning of the learning rate schedule. Binary cross-entropy, defined as $L = -[y \cdot \log(p) + (1-y) \cdot \log(1-p)]$, is the theoretically correct loss for a single-sigmoid binary classifier.

Training ran for up to 50 epochs with batch size 32, and the training partition was reshuffled at the start of each epoch. Two callbacks governed the dynamics: ReduceLROnPlateau cut the learning rate by a factor of 0.2 whenever validation loss showed no improvement for five consecutive epochs, and EarlyStopping halted training if validation accuracy did not improve over ten consecutive epochs, restoring the best checkpoint. All experiments ran in Python 3.8 with TensorFlow 2.x, Keras, and OpenCV 4.5 on an NVIDIA GPU-equipped workstation.

IV. Results and Discussion

A. Classification Performance on the Test Set

The trained model was evaluated on the 600-image held-out test partition (300 per class), which played no role in training or hyperparameter selection. Four metrics were computed: overall accuracy, class-wise precision, class-wise recall, and class-wise F1-score. Precision measures what fraction of a class's predicted positives are genuine; recall measures what fraction of a class's actual positives the model catches; F1-score is the harmonic mean of the two and captures both aspects of quality in a single number. Full results are presented in Table II below.

II. Classification Performance on the Held-Out Test Set

Class	Precision	Recall	F1-Score	Support
Naturally Ripened	0.94	0.92	0.93	300
Chemically Ripened	0.92	0.94	0.93	300
Overall Accuracy	0.93	—	—	600

The model reached 93 percent overall accuracy. Precision of 0.94 for naturally ripened and 0.92 for chemically ripened indicates that the overwhelming majority of the model's positive predictions in each class are correct. Recall of 0.92 and 0.94 respectively shows that almost all true members of each class are correctly identified. The slightly higher recall for the chemically ripened class (0.94 against 0.92) is the preferable bias in a food safety setting: the model is marginally more likely to flag a treated banana than to miss one, which is the direction that reduces potential harm. F1-scores of 0.93 for both classes confirm that this performance is consistent and balanced rather than driven by a bias toward one category.

B. Learning Curve Behaviour

Training accuracy rose sharply from about 60 percent to 85 percent during the first fifteen epochs as the early convolutional layers picked up discriminative low-level features. Validation accuracy tracked closely throughout, confirming

 ©IJSRED: All Rights are Reserved

genuine generalisation rather than memorisation of training examples. Between epochs 16 and 35 both metrics continued to improve at a slower pace, with the ReduceLROnPlateau callback triggering twice to reduce the learning rate from 0.001 down to 0.0002 and then to 0.00004, enabling finer-grained parameter adjustment. The training-validation loss gap stayed small throughout this phase, at roughly 0.04 to 0.06. In the final phase, both curves plateaued: training accuracy near 96 percent and validation near 93 percent. The EarlyStopping callback did not activate, indicating the model was still making incremental gains up to epoch 50.

C. Practical Advantages of This Approach

Four properties make the system attractive for real deployment. First, it is entirely non-destructive: the assessed banana remains intact and saleable after inspection because only a surface photograph is needed. Second, inference takes 15 to 30 milliseconds per image on GPU hardware and under 200 milliseconds on CPU-only hardware, making real-time operation practical on an inspection line or at a market stall. Third, the only hardware required is a standard smartphone camera, putting the system within reach of low-resource settings including rural markets and small distribution centres. Fourth, the augmentation strategy applied during training — combining rotation, flipping, zoom, brightness variation, and spatial shifts — gives the model reasonable robustness to the incidental differences in lighting, angle, and framing that arise in real capture conditions.

D. Limitations and Error Analysis

Several limitations need to be stated clearly. The training dataset of 4,000 images is modest by the standards of production-grade classifiers; expanding to tens of thousands of images, and covering a wider range of banana cultivars, geographic origins, and storage durations, would be expected to improve both accuracy and robustness. Manual review of the misclassified test images showed that the great majority involved either extreme lighting conditions or bananas at transitional ripeness stages where visual markers of both classes overlapped. The chemically ripened class recall of 0.94 means roughly 6 in every 100 treated bananas are passed as natural; in a safety-critical deployment this false-negative rate would need to be reduced further. The system also addresses only the binary ripening-method question and says nothing about bruising severity, fungal infection, ripeness stage on a 1-to-7 scale, or remaining shelf life.

V. Conclusion

This paper described a CNN-based image classifier for distinguishing naturally ripened from chemically ripened bananas. The work was motivated by the real public health hazard posed by calcium carbide treatment, which deposits arsenic and phosphorus compounds on fruit eaten daily by large numbers of people. A four-thousand-image balanced dataset was assembled from retail markets and laboratory settings, preprocessed through resizing, pixel normalisation, and real-time augmentation, and used to train a purpose-designed six-layer sequential CNN. On a held-out test set the model reached 93 percent overall accuracy with balanced F1-scores of 0.93 for both classes and a safety-favourable recall of 0.94 for the chemically ripened class.

Several lines of future work are planned. The training dataset will be expanded to cover multiple banana cultivars — Cavendish, Robusta, Nendran, Poovan — and a wider range of market environments and imaging conditions. Transfer learning using MobileNetV2, EfficientNet-B0, and DenseNet-121 will be evaluated as alternatives to the from-scratch architecture, since these pre-trained backbones may deliver higher accuracy at lower training cost. A TensorFlow Lite mobile application will be developed to bring the detection capability to consumers directly at the point of purchase. Multimodal approaches combining visible-spectrum images with compact near-infrared sensors or low-cost electronic nose modules will be explored as a route to accuracy beyond what visible imagery alone can support. Finally the framework will be extended to cover other fruits commonly treated with calcium carbide or ethephon, including mangoes, tomatoes, papayas, and grapes, to broaden the scope and public health impact of the tool.

Acknowledgment

 ©IJSRED: All Rights are Reserved

The authors thank the Department of Computer Science and Engineering (Artificial Intelligence) at Ballari Institute of Technology and Management for access to computing infrastructure and laboratory facilities. They are also grateful to their project supervisor for ongoing technical guidance, and to the fruit vendors and farm operators in Ballari whose cooperation made real-market dataset collection possible. The food safety officer and agricultural scientist who independently annotated the images are sincerely acknowledged for their time and expertise.

References

[1] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, article 1419, pp. 1–10, Sep. 2016.

[2] I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, and C. McCool, "DeepFruits: A fruit detection system using deep neural networks," Sensors, vol. 16, no. 8, article 1222, pp. 1–23, Aug. 2016.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, Jun. 2017.

[4] M. T. Islam, M. A. Wahid, and M. S. Uddin, "Identification of chemically ripened fruits using image processing and machine learning techniques," in Proc. IEEE Int. Conf. on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, Bangladesh, 2019, pp. 1–5.

[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.

[6] M. Bhargava and A. Bhargava, "Fruits freshness grading using deep learning," Int. Journal of Engineering Research and Technology, vol. 9, no. 6, pp. 1–6, Jun. 2020.

[7] P. Sirimanna, R. Karunanayaka, and H. Warnapura, "Detection of calcium carbide ripened mangoes using image processing techniques," in Proc. Int. Research Conf. on Smart Computing and Systems Engineering (SCSE), Colombo, Sri Lanka, 2021, pp. 72–77.

[8] F. Chollet, Deep Learning with Python. Shelter Island, NY, USA: Manning Publications, 2018.

[9] World Health Organization (WHO), "Food safety: Chemical hazards in fruit and vegetables," WHO Technical Report Series, Geneva, Switzerland, 2020.

[10] M. A. Hossen, M. M. Hasan, and S. M. Masud, "A review on machine learning approaches for fruit ripeness classification and quality assessment," Journal of Food Engineering, vol. 289, article 110175, pp. 1–14, Jan. 2021.

[11]] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. Int. Conf. on Learning Representations (ICLR), San Diego, CA, USA, 2015

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.

 ©IJSRED: All Rights are Reserved