(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

An Integrated Deep Learning and Computer Vision Framework for Non-Destructive Assessment of Apple Nutritional Quality

Amisha Singh[1], Nishant Singh [2], Ankit Kumar[3], Uttam Sharma[4], Basant Kumar Das[5], Harsh P. Sharma[6*], Anuj Yadav[7]

[1] Department of Applied Sciences & Humanities [Mathematics], Indira Gandhi Delhi Technical University for Women, New Delhi 110006, India

[2] Department of Mechanical Engineering, College of Technology, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut-250110, Uttar Pradesh, India

[3] Department of Process and Food Engineering, College of Technology, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut-250110, Uttar Pradesh, India

[4] Department of Information Technology, College of Technology, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut-250110, Uttar Pradesh, India

[5]Department of Mechanical Engineering, Arka Jain University, Jamshedpur-832108, Jharkhand, India

[6]Department of Food Process Technology, College of Post-Harvest Technology and Food Processing, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut-250110, Uttar Pradesh, India

[7]Department of Agriculture Keral Verma Subharti College of Science Swami Vivekanand Subharti University, Meerut

*Corresponding author: harshsharma1983@gmail.com

DOI:https://doi.org/10.63001/tbs.2026.v21.i02.S.I(2).pp21723-21756

KEYWORDS

Artificial intelligence;
Deep learning;
Vision Transformer;
EfficientNet;
Explainable artificial intelligence;
Apple quality;
Nutritional prediction.

Article Info

Received: 02.01.2026

Revised : **12-03-2026**

Accepted: 15.04.2026

Published:26.04.2026

Abstract

Rapid and non-destructive assessment of apple nutritional quality is essential for intelligent postharvest management and automated grading. This study proposes an integrated framework combining RGB computer vision, transfer learning, explainable artificial intelligence (XAI), and a decision support system (DSS) for simultaneous prediction of multiple nutritional attributes, including soluble solids content (SSC), moisture, firmness, vitamin C, total phenolic content (TPC), titratable acidity, total sugars, and dry matter. A dataset comprising 720 Red Delicious apples represented by 4,800 multi-view RGB images was used to develop and evaluate five deep learning architectures: CNN, ResNet50, DenseNet121, EfficientNet-B3, and Vision Transformer (ViT). Among the evaluated models, ViT achieved the highest overall performance with $R^2 = 0.964$, RMSE = 0.25, MAE = 0.18, MAPE = 2.10%, Pearson's r = 0.982, and NSE = 0.960, while attribute-specific prediction accuracies reached $R^2 = 0.979$ for SSC, 0.957 for moisture, 0.962 for firmness, 0.952 for vitamin C, and 0.957 for TPC. Bland–Altman analysis demonstrated excellent agreement with laboratory measurements, with minimal mean bias (SSC: 0.03, moisture: −0.04, firmness: −0.21, and vitamin C: 0.08). Grad-CAM and SHAP analyses confirmed that the models relied on physiologically meaningful features, with peel colour and texture contributing most strongly to nutritional prediction. The proposed DSS achieved an overall grading accuracy of 96.1% (692/720 samples). Although ViT provided the highest prediction accuracy, EfficientNet-B3 achieved comparable performance ($R^2 = 0.958$) with substantially lower computational requirements (48 MB model size; 21 ms/image; 48 FPS), making it more suitable for real-time edge deployment. These findings demonstrate that the proposed framework provides an accurate, rapid, and scalable solution for non-destructive nutritional assessment and automated apple quality grading.

Nomenclature

Roman Symbols

Symbol	Description	Unit
F	Fruit firmness	N
I	Original pixel intensity	–
I_{max}	Maximum pixel intensity	–
I_{min}	Minimum pixel intensity	–
I_{norm}	Normalized pixel intensity	–
L	Loss function	–
n	Number of observations (samples)	–
N	Total number of samples	–
R^2	Coefficient of determination	–
r	Pearson correlation coefficient	–

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

Symbol	Description	Unit
x	Input RGB image	–
y	Measured (ground truth) nutritional value	–
$\hat{y}$	Predicted nutritional value	–

Greek Symbols

Symbol	Description
α	MixUp/CutMix interpolation parameter
σ	Standard deviation; Gaussian blur parameter
ω	Importance weight of the *k*th feature map in Grad-CAM

Performance Metrics

Symbol	Description
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
NSE	Nash–Sutcliffe Efficiency
PSNR	Peak Signal-to-Noise Ratio
RMSE	Root Mean Square Error
SSIM	Structural Similarity Index Measure
FPS	Frames Per Second

Nutritional Parameters

Symbol	Description	Unit
SSC	Soluble Solids Content	°Brix
MC	Moisture Content	%
F	Firmness	N
TA	Titratable Acidity	% malic acid
VC	Vitamin C	mg 100 g^{-1}
TS	Total Soluble Sugars	g 100 g^{-1}
TPC	Total Phenolic Content	mg GAE 100 g^{-1}
DM	Dry Matter	%

Machine Learning and Computer Vision Abbreviations

Abbreviation	Full Form
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
ResNet50	Residual Network-50
DenseNet121	Densely Connected Convolutional Network-121
EfficientNet-B3	EfficientNet-B3
ViT-B16	Vision Transformer-Base/16
RGB	Red–Green–Blue
XAI	Explainable Artificial Intelligence
Grad-CAM	Gradient-weighted Class Activation Mapping
SHAP	Shapley Additive explanations
DSS	Decision Support System
CLAHE	Contrast Limited Adaptive Histogram Equalization

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

Abbreviation	Full Form
HDF5	Hierarchical Data Format Version 5
ImageNet	Large-scale Image Database for Transfer Learning

1.0 Introduction

Apple (*Malus domestica* Borkh.) is a globally cultivated temperate fruit crop with major nutritional and commercial value. Apples are described as important fruits in global markets and are valued for both market quality and nutritional quality [1], [2]. They are also widely appreciated for flavour and health-promoting compounds, and several studies frame apples as foods associated with reduced risk of chronic and age-related disorders [3], [4].

The nutritional quality of apples depends on a broad set of physicochemical and bioactive constituents. Studies across apple germplasm and cultivar evaluations identify soluble solids content (SSC), titratable acidity, firmness, vitamin C, sugars, phenolics, flavonoids, and antioxidant capacity as central quality indicators linked to maturity, sensory performance, and nutritional value [3], [5]. Considerable biochemical variation exists among apple genotypes, including differences in polyphenols, tannins, flavonoids, anthocyanins, sugars, and carotenoids, which supports the importance of varietal characterization in quality evaluation and breeding [6]. Climate and genetic origin also shape apple composition, with accessions and yearly climate features acting as major determinants of metabolite profiles and fruit characteristics [3].

Bioactive compounds are especially important because they contribute to both nutritional quality and functional value. Reviews and genotype studies emphasize the relevance of phenolic compounds to apple appearance, pulp density, antioxidant protection, and nutritional potential [7]. Specific compounds such as chlorogenic acid and catechin were quantified at substantial levels in the endemic "Long Apple" variety, alongside high vitamin C, total sugar, malic acid, firmness, soluble solids, and antioxidant capacity, illustrating how cultivar-specific biochemical profiles can elevate commercial and health value [5]. Local apple accessions also tended to show higher antioxidant-related traits than foreign and commercial ones, indicating that breeding and germplasm selection remain important routes for improving nutraceutical quality [3].

Maintaining these attributes after harvest remains difficult because apple quality changes continuously during storage and handling. Apples are highly perishable and require proper preservation to reduce degradation of macro- and micronutrients and to extend shelf life [4]. Storage studies show that internal attributes such as firmness, pH, moisture content, and SSC are monitored over extended storage periods because they shift measurably with postharvest conditions and biological stressors [8]. Bruising and other handling damage are also relevant because they degrade nutritional value, increase susceptibility to decay, and alter quality and safety status during postharvest operations [9], [10].

For this reason, rapid assessment of internal quality has become a priority. Laboratory-based evaluation remains the conventional standard, but it usually requires destructive sample preparation and specialized operation [1]. Common reference methods include refractometry for SSC, HPLC or titrimetric approaches for vitamin C, and Bradford-type colorimetric assays for soluble protein, but these methods are labor-intensive and unsuitable for continuous large-scale monitoring [2]. Reviews of food-quality sensing similarly conclude that destructive physicochemical analysis limits rapid and scalable quality control, especially when repeated monitoring is needed across storage and marketing chains [10], [11].

Non-destructive optical sensing has therefore become a major research direction for apple quality evaluation. Hyperspectral imaging is repeatedly described as a promising approach because it combines spectral information with spatial information, enabling simultaneous assessment of external appearance and internal composition [10], [12]. In apples, hyperspectral methods have been used to predict SSC, firmness, moisture content, pH, starch, vitamin C, and titratable acidity, showing that multiple nutritional and physicochemical traits can be estimated non-destructively from reflectance data [1], [8], [11]. Some studies also show that feature wavelength selection can reduce redundancy and computational burden while preserving or improving predictive performance, which is important for real-time deployment [4], [9].

Model performance, however, varies by target attribute and sensing design. Near-infrared hyperspectral reflectance imaging predicted SSC and moisture content accurately in stored Fuji apples, but all tested models failed to predict firmness well in one study, illustrating that not all internal traits are equally inferable from the same spectral setup [8]. Other work reported stronger firmness prediction, with ANN achieving an R^2 of 0.910 and decision tree models also performing well for firmness, while SSC prediction was more moderate and model-dependent [13]. More recent rotational hyperspectral imaging studies expanded prediction to four nutritional qualities—starch, vitamin C, SSC, and titratable acid—and reported prediction coefficients above 0.90 for all four optimal models [1].

Artificial intelligence has accelerated this transition from laboratory assays to automated quality inference. Reviews of hyperspectral food analysis conclude that machine learning enables rapid, high-accuracy classification and regression, while deep learning is particularly promising because of its feature-learning capacity for real-time applications . Across postharvest fruit applications, deep learning systems have been used for automated defect detection, disease grading, ripeness classification, and sorting, often to reduce labor dependence and postharvest losses [14], [15]. Convolutional models can extract color, shape, texture, and surface features directly from images, reducing the need for manual feature engineering [16], [17].

In apple quality research specifically, machine learning and deep learning have been applied to internal-quality prediction from hyperspectral and near-infrared data with strong results. Deep learning combined with hyperspectral imaging predicted Fuji apple SSC with $Rp^2 = 0.9436$ and low error, supporting its feasibility for rapid non-destructive assessment [18]. For intact apples, diffuse reflectance spectra coupled with PLSR or CNN also produced high SSC prediction performance, with best models reaching $Rp = 0.96$ for PLSR and $Rp = 0.95$ for CNN [19]. Deep learning-enabled hyperspectral modeling across six apple varieties further showed robust multi-variety prediction of vitamin C, SSC, and soluble protein, including cross-year validation, which is important for generalization beyond single-cultivar datasets [2].

At the same time, practical deployment still faces several limitations. Existing studies often concentrate on one or two traits, most commonly SSC and firmness, whereas integrated nutritional assessment requires simultaneous prediction of multiple interacting indicators [1]. Many systems also rely on controlled illumination, standardized region-of-interest extraction, and computationally heavy hyperspectral pipelines, which can complicate industrial implementation despite high laboratory accuracy [12]. Seasonal variation and broader environmental variability remain open challenges for model robustness, and review work identifies the need for further research to incorporate such variation into real-time hyperspectral learning systems [10]

These gaps help explain interest in lower-cost and more deployable imaging strategies. RGB imaging is useful for capturing external traits such as size, color, and texture, while hyperspectral or near-infrared methods are more closely linked to internal quality through spectral signatures and reflectance behavior [17]. Hybrid systems that combine RGB with spectral imaging have achieved high classification accuracy in other fruits and are being positioned for real-time sorting and noninvasive grading on postharvest lines. Likewise, lightweight embedded models for agricultural grading demonstrate that real-time AI deployment is technically achievable when model efficiency is considered alongside accuracy [15].

Overall, the literature supports a clear transition from destructive chemical assays toward AI-assisted non-destructive sensing for apple nutritional evaluation. Apples show substantial biochemical variability driven by genotype, climate, and postharvest conditions, and the most informative quality traits include SSC, acidity, firmness, vitamin C, moisture-related properties, sugars, and phenolic compounds [3], [6]. Hyperspectral imaging, near-infrared sensing, and learning-based models now provide strong evidence that internal quality can be estimated rapidly and non-destructively, but wider commercial adoption still depends on multi-attribute prediction, stronger generalization under field and storage variability, reduced system cost, and more interpretable AI models [8], [12].

2. Materials and Methods

2.1 Overall Research Framework

The present study proposes an integrated artificial intelligence framework for non-destructive assessment of apple nutritional quality by combining RGB computer vision, transfer learning, explainable artificial intelligence (XAI), and intelligent decision-support techniques. Unlike conventional laboratory methods, which are accurate but destructive, time-consuming, and unsuitable for large-scale commercial applications, the proposed framework enables rapid prediction of internal nutritional attributes from RGB images [20], [21]. The framework establishes relationships between external visual characteristics and internal nutritional properties, building on previous studies that demonstrated the effectiveness of RGB imaging, hyperspectral imaging, and Vis/NIR spectroscopy for predicting SSC, firmness, acidity, moisture, dry matter, vitamin C, phenolic compounds, and fruit ripeness [22], [23]. The methodology comprises seven stages: sample collection and nutritional characterization, RGB image acquisition, image preprocessing, transfer learning-based model development, nutritional attribute prediction, XAI-based model interpretation, and intelligent decision support for automated grading and precision horticulture [24], [25].

The proposed framework simultaneously predicts SSC, moisture content, firmness, titratable acidity, vitamin C, total soluble sugars, dry matter, and total phenolic content, which are key indicators of sweetness, texture, flavor, nutritional value, storability, antioxidant capacity, and consumer acceptance. Consequently, the framework provides a scalable solution for intelligent quality grading, postharvest management, and precision horticulture. Figure 1 presents the overall research framework adopted in this study [26].

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

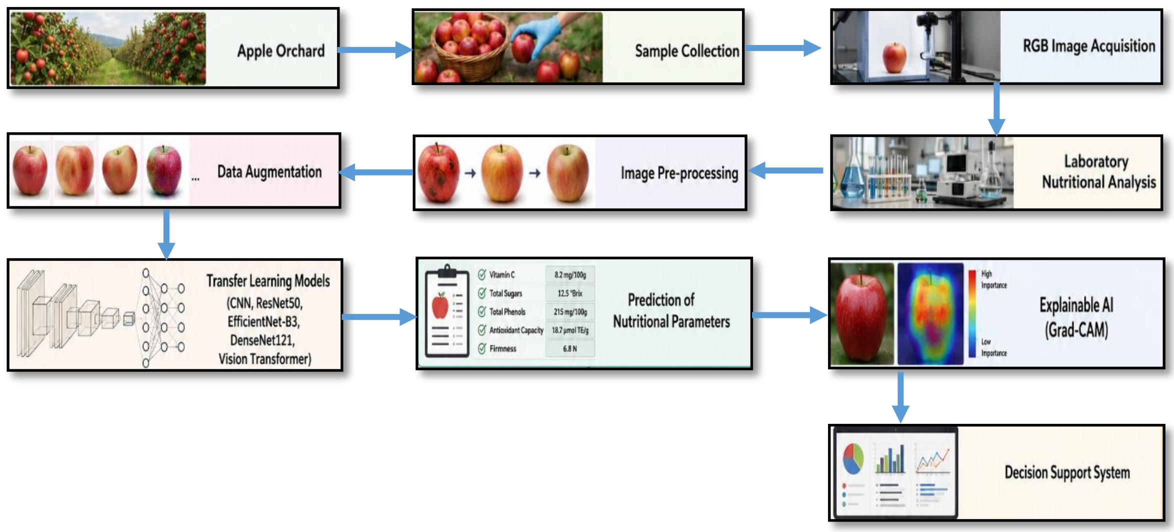


Figure 1. Overall research framework

2.2 Apple Sample Collection

Apple (*Malus domestica* Borkh.) fruits of the commercially important 'Red Delicious' cultivar were collected during the commercial harvest season from six representative orchards. Fruits were harvested at commercial maturity based on standard maturity indices, including peel colour, starch degradation, firmness, and soluble solids content (SSC) [27], [28]. Only healthy fruits free from insect damage, mechanical injury, and physiological disorders were selected to ensure sample uniformity for image-based quality analysis [29], [30].

A total of 720 apples, representing different maturity stages and commercial quality grades, were collected using a completely randomized sampling design to capture natural variability associated with orchard location, canopy position, light exposure, irrigation, and nutritional management [31]. Following harvest, the samples were transported to the laboratory in insulated containers and conditioned for 24 h at 20 ± 2 °C and 65 ± 5% relative humidity before RGB image acquisition and nutritional analysis. Each fruit was assigned a unique QR-coded identification number to maintain traceability between image data and reference laboratory measurements throughout the experimental workflow, thereby ensuring reliable training and validation of the deep learning models [32]. The experimental characteristics of the collected apple samples are summarized in Table 1.

Table 1: Experimental characteristics of apple samples

S. No.	Parameter	Description
1	Apple cultivar	Red Delicious
2	Total samples	720 fruits
3	Orchard locations	6 commercial orchards
4	Harvest maturity	Commercial maturity
5	Experimental design	Completely randomized
6	Storage before imaging	24 h
7	Laboratory temperature	20 ± 2 °C
8	Relative humidity	65 ± 5%
9	Sample identification	QR-coded unique ID

2.3 Ground Truth Nutritional Analysis

Ground truth nutritional measurements were performed immediately after RGB image acquisition to generate reference labels for supervised model training and validation using standard laboratory methods for fruit quality evaluation. The evaluated nutritional parameters and their corresponding analytical methods are summarized in Table 2 [26], [33]. The evaluated attributes included soluble solids content (SSC), moisture content, firmness, titratable acidity (TA), vitamin C (VC), total soluble sugars, total phenolic content (TPC), and dry matter, which are widely recognized as key indicators of apple quality, maturity, nutritional value, and consumer acceptability [34], [35].

SSC was measured using a digital refractometer, firmness with a texture analyzer, TA by NaOH titration (expressed as malic acid), moisture content and dry matter by the gravimetric oven-drying method, vitamin C using HPLC, and TPC by the Folin–Ciocalteu assay, following established analytical protocols [36], [37]. The resulting laboratory measurements were linked to their corresponding RGB images through unique sample identifiers and used as ground truth labels for multi-attribute regression.

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

Table 2: Ground truth nutritional parameters

S. No.	Parameter	Unit	Instrument
1	Soluble Solids Content	°Brix	Digital refractometer
2	Moisture Content	%	Hot air oven
3	Firmness	N	Texture analyzer
4	Titratable Acidity	% malic acid	Automatic titrator
5	Vitamin C	mg/100 g	HPLC
6	Total Sugars	g/100 g	UV–Visible spectrophotometer
7	Total Phenolic Content	mg GAE/100 g	Folin–Ciocalteu method
8	Dry Matter	%	Oven drying

2.3.1 Soluble Solids Content: SSC was measured using a calibrated digital refractometer with an accuracy of ±0.1 °Brix. Fresh juice extracted from the equatorial region of each apple was filtered and analyzed at room temperature [34].

2.3.2 Moisture Content: Moisture content was determined gravimetrically using the hot-air oven method at 105°C until constant weight was achieved. Moisture percentage was calculated as [35]:

$$MC(\%) = \frac{W_i - W_d}{W_i} \times 100$$

Where W_i is the initial sample weight (g), and W_d is the dry weight (g).

2.3.3 Dry Matter: Dry matter content was calculated as [35]:

$$DM(\%) = \frac{W_d}{W_i} \times 100$$

2.3.4 Titratable Acidity: Titratable acidity was determined by titrating apple juice against 0.1 N NaOH using phenolphthalein as an indicator and expressed as percentage malic acid [34], [35].

2.3.5 Vitamin C: Vitamin C concentration was quantified using High-Performance Liquid Chromatography (HPLC) equipped with a UV detector operating at 254 nm. Standard calibration curves were prepared using analytical-grade L-ascorbic acid [34], [35].

2.3.6 Total Phenolic Content: Total phenolics were measured according to the Folin–Ciocalteu colorimetric assay and expressed as milligrams of gallic acid equivalents (GAE) per 100 g fresh weight [34], [35].

2.3.7 Firmness: Fruit firmness was measured using a digital texture analyzer equipped with an 11-mm stainless steel probe at a penetration speed of 5 mm s^{-1}. Measurements were recorded at two opposite equatorial positions, and the mean value was used for analysis. The descriptive statistics of fruit firmness are presented in Table 3 [36], [37].

Table 3: Descriptive statistics of nutritional parameters

S. No.	Parameter	Minimum	Maximum	Mean ± SD
1	SSC (°Brix)	10.4	17.8	13.8 ± 1.6
2	Moisture (%)	80.2	87.9	84.6 ± 1.8
3	Firmness (N)	42.8	81.5	63.9 ± 8.4
4	Titratable Acidity (%)	0.18	0.72	0.43 ± 0.11
5	Vitamin C (mg/100 g)	3.5	10.8	6.9 ± 1.4
6	Total Sugars (g/100 g)	8.2	14.3	11.1 ± 1.2
7	Total Phenolics (mg GAE/100 g)	95	248	168 ± 32
8	Dry Matter (%)	12.1	19.8	15.4 ± 1.5

The measured nutritional values served as the ground truth labels for supervised training of the deep learning models. Each RGB image was linked to its corresponding laboratory measurements through the unique sample identification code, enabling regression-based prediction of nutritional parameters using image-derived features.

2.4 RGB Image Acquisition

A standardized RGB imaging system, illustrated in Figure 2, was developed to acquire high-resolution images of apple fruits under controlled conditions for deep learning-based prediction of nutritional attributes. Images were captured inside a matte-black imaging chamber to minimize ambient light interference, background noise, and reflection artifacts. Uniform illumination was provided by four dimmable LED panels positioned at 45°, maintaining an intensity of 1200 ± 25 lux to ensure consistent lighting and reduce shadows and specular reflections [38], [39]. A Canon EOS 90D camera equipped with a 50 mm f/1.8 lens was mounted 500 mm above the sample platform and operated in manual mode (f/8, 1/125 s, ISO 100) to maintain fixed imaging geometry. Images were acquired at a resolution of 6960 × 4640 pixels, stored in RAW (.CR3) and TIFF formats, and colour-calibrated using an X-Rite ColorChecker Passport to ensure consistent colour reproduction [40].

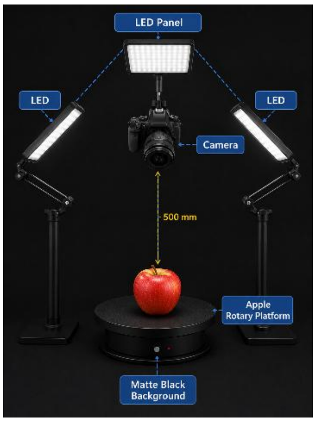


Figure 2. RGB image acquisition system

Each apple was placed on a motorized rotary platform, and four RGB images were captured at 90° rotational intervals to obtain complete surface coverage. The final dataset comprised 4,800 RGB images from 720 apple samples, with each image linked to a unique sample identifier to maintain traceability with the corresponding laboratory-derived nutritional measurements. The RGB image acquisition specifications are shown in Table 4.

Table 4: RGB image acquisition specifications

S. No.	Parameter	Specification
1	Camera	Canon EOS 90D DSLR
2	Sensor resolution	32.5 MP APS-C CMOS
3	Lens	50 mm f/1.8 prime lens
4	Image resolution	6960 × 4640 pixels
5	Image format	RAW (.CR3), TIFF
6	Aperture	f/8
7	Shutter speed	1/125 s
8	ISO	100
9	Camera distance	500 mm
10	LED panels	4
11	Colour temperature	5600 K
12	Illumination intensity	1200 ± 25 lux
13	Background	Matte black acrylic
14	Images per fruit	4
15	Total RGB images	4,800

2.5 Image Pre-processing

The acquired RGB images were pre-processed to enhance image quality, remove background noise, normalize illumination, and generate standardized inputs for deep learning models. The pre-processing pipeline was implemented using OpenCV 4.9, Python 3.11, and scikit-image. The overall workflow is illustrated in Figure 3, while the corresponding image pre-processing parameters are summarized in Table 5.

Initially, RAW images were converted to 24-bit RGB format and resized to 512 × 512 pixels using bicubic interpolation. Background segmentation was performed using adaptive Otsu thresholding followed by morphological operations and connected-component analysis to isolate the fruit region. The segmented apple was automatically cropped using its minimum bounding rectangle [41].

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

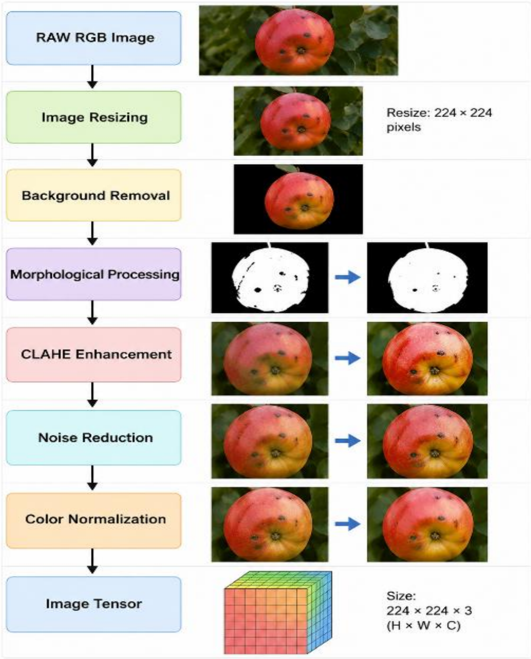


Figure 3. Image pre-processing workflow

Table 5: Image preprocessing parameters

S. No.	Operation	Method	Parameter
1	Image resizing	Bicubic interpolation	512 × 512 pixels
2	Segmentation	Otsu thresholding	Adaptive
3	Morphology	Opening & Closing	5 × 5 kernel
4	Contrast enhancement	CLAHE	Clip limit = 2.0
5	Noise reduction	Median filter	3 × 3 kernel
6	Edge preservation	Bilateral filter	σ = 75
7	Normalization	Min–Max scaling	[0,1]
8	Color correction	Histogram matching	Reference image

To reduce illumination variability, Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied to the luminance channel in the CIELAB colour space, followed by histogram-based colour normalization using a ColorChecker reference image. Image noise was further reduced using a 3 × 3 median filter and bilateral filtering while preserving edge information. Finally, pixel intensities were normalized to the range [0,1] using [42]:

$$I_{\text{norm}} = \frac{I - I_{\text{min}}}{I_{\text{max}} - I_{\text{min}}}$$

Where (I) is the original pixel intensity, and I_{min} and I_{max} denote the minimum and maximum intensity values, respectively. The processed images were then converted into tensor format and stored in the HDF5 database for efficient loading during deep learning model training.

2.6 Data Augmentation

To improve the generalization capability of the deep learning models and reduce overfitting, an extensive data augmentation strategy was applied to the training images. The augmentation process introduced realistic variations in fruit orientation, appearance, and imaging conditions while preserving the original sample labels [43], [44]. The adopted data augmentation techniques and their corresponding parameters are summarized in Table 6. Each training image was randomly transformed using a combination of geometric and photometric operations, including horizontal and vertical flipping, rotation, translation, zooming, random cropping, brightness adjustment, contrast variation, and Gaussian blur, which are widely used augmentation methods for improving invariance and reducing overfitting in CNN-based image classification [45], [46]. Flipping and random cropping increase orientation and spatial variation, while brightness and contrast perturbations help the model tolerate imaging variability rather than memorizing fixed visual patterns [47]. Advanced mixed-sample augmentation was also incorporated through MixUp and CutMix, where MixUp generates synthetic examples by convex interpolation of image pairs and labels, and CutMix replaces local image regions with patches from another sample while mixing labels proportionally to patch area [48]. These two-image augmentation methods tend to improve robustness and generalization beyond conventional single-image transformations, and CutMix in particular improves resistance to input corruption while MixUp reduces memorization and encourages smoother decision boundaries [49]. This augmentation pipeline increased the effective training set from 3,360 original images to about 20,000 samples, thereby substantially increasing training diversity, which is the central mechanism by which augmentation improves deep learning performance under limited data conditions [50].

Table 6: Data augmentation strategy

S. No.	Augmentation	Parameter
1	Rotation	±30°
2	Horizontal flip	Probability = 0.5
3	Vertical flip	Probability = 0.5
4	Translation	±15%
5	Scaling	0.8–1.2
6	Brightness	±20%
7	Contrast	±20%
8	Gaussian blur	σ = 0.5–1.5
9	MixUp	α = 0.2
10	CutMix	α = 1.0

2.7 Dataset Preparation

The complete RGB image dataset was randomly divided into mutually exclusive training, validation, and testing subsets, ensuring that all images from the same apple sample were assigned to a single subset to prevent data leakage and overestimated model performance [51], [52]. This sample-level partitioning improves the reliability and generalization of the deep learning models by preventing information leakage between datasets [53]. The dataset distribution is summarized in Table 7. The dataset was partitioned into 70% training, 15% validation, and 15% testing, corresponding to 3,360, 720, and 720 images, respectively, which is consistent with common three-way train/validation/test designs used in deep learning workflows [54]. The training subset was used to optimize model parameters, whereas the validation subset was used for hyperparameter tuning and early stopping during model development . The independent test subset was kept unseen throughout training and tuning and was used exclusively for final performance evaluation, as recommended to obtain an unbiased estimate of model generalization [55].

Table 7: Dataset partitioning

S. No.	Dataset	Images	Percentage
1	Training	3,360	70%
2	Validation	720	15%
3	Testing	720	15%
4	Total	4,800	100%

2.8 Deep Learning Model Development

To relate external visual features to internal nutritional attributes, five deep learning models were developed using transfer learning: a custom CNN, ResNet-50, DenseNet121, EfficientNet-B3, and ViT-B16, reflecting architectures widely used in fruit and agricultural image analysis [56]. ResNet-50, DenseNet121, EfficientNet-B3, and Vision Transformers were selected because they have shown strong performance for fruit quality, disease, and freshness analysis, with ResNet extracting hierarchical image features efficiently, DenseNet improving feature reuse, EfficientNet balancing accuracy and computational cost, and ViT performing competitively in fruit imaging tasks [57].

All models were implemented with ImageNet-pretrained weights, since transfer learning reuses features learned from large benchmark datasets and commonly improves convergence, training efficiency, and performance when agricultural image datasets are limited [58]. In fruit imaging, pretrained models are routinely adapted by replacing the original classification layer with a task-specific head, often including global average pooling and dense layers, to better match the target application [59].

Because nutritional traits are continuous variables, the original classifier was replaced with a regression head composed of global average pooling, fully connected layers with batch normalization, ReLU, and dropout, followed by an output layer in which each neuron predicted one nutritional parameter [60]. This design supports simultaneous prediction of multiple nutritional attributes from RGB images while retaining the feature-extraction strength of pretrained vision backbones [61].

The overall deep learning framework and the architectures of the evaluated models are illustrated in Figure 4, while the architectural details and model configurations are summarized in Table 8.

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

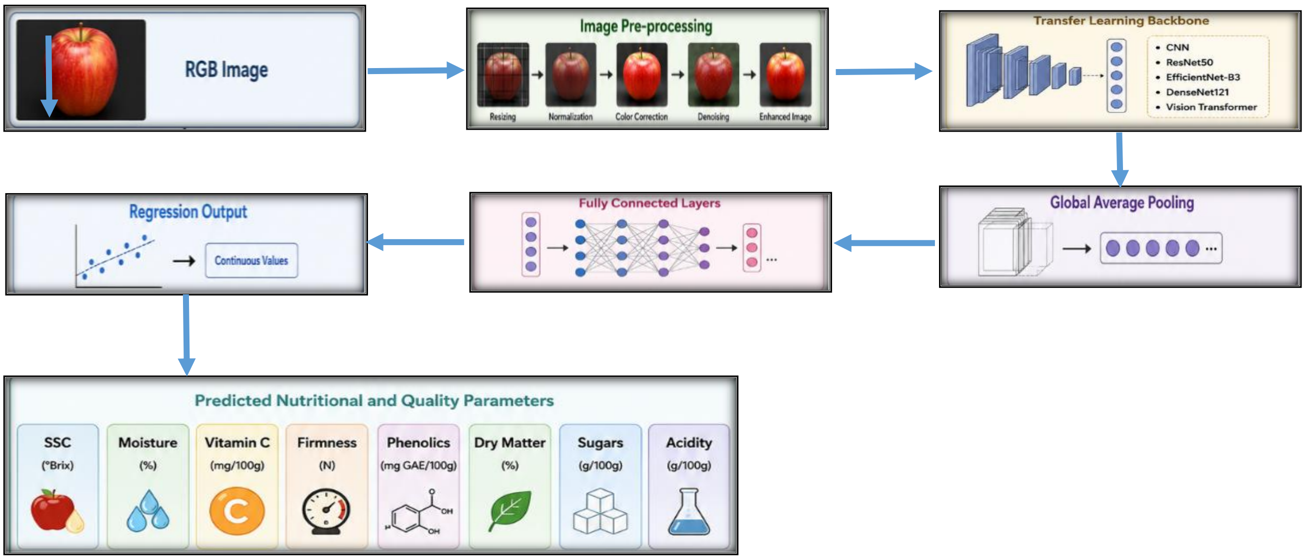


Figure 4. Deep learning framework

Table 8: Deep learning architectures evaluated

S. No.	Model	Parameters (Million)	Image Size	Transfer Learning	Task
1	CNN	5.3	512×512	No	Regression
2	ResNet50	25.6	512×512	Yes	Regression
3	EfficientNet-B3	12.0	512×512	Yes	Regression
4	DenseNet121	8.0	512×512	Yes	Regression
5	Vision Transformer	86.4	512×512	Yes	Regression

2.9 Transfer Learning Strategy

Transfer learning was used to address the limited availability of labeled nutritional images and to improve training efficiency, since deep learning models often underperform with scarce labels and benefit from reusing representations learned on large source datasets [62]. Accordingly, all models were initialized with ImageNet-pretrained weights and fine-tuned on the apple RGB dataset, consistent with common practice in agricultural and fruit-image analysis [63].

In the first stage, the feature-extraction backbone was kept frozen and only the regression head was optimized, a standard transfer-learning setup that preserves generic pretrained visual features while reducing computational cost and data requirements [64]. In the second stage, the final 30% of backbone layers were unfrozen and jointly fine-tuned with a lower learning rate, allowing the pretrained features to adapt to apple nutritional prediction while helping control overfitting during domain-specific training [65].

Model optimization was performed by minimizing the Mean Squared Error (MSE) loss function:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

Where

- y_i = measured nutritional value
- $\hat{y}_i$ = predicted value
- N = total samples

To improve robustness, the overall optimization objective combined MSE loss with L2 regularization:

$$L = L_{MSE} + \lambda \|W\|_2^2$$

Where

$$\lambda = 0.0001.$$

2.10 Hyperparameter Optimization

Hyperparameter optimization was performed to identify the optimal training configuration and maximize model performance. A Bayesian optimization strategy combined with five-fold cross-validation was employed to efficiently explore the hyperparameter space and improve model generalization [66]. The optimized hyperparameters included the learning rate, batch size, weight decay, dropout probability, optimizer, number of training epochs, learning-rate scheduler, and early stopping patience, all of which significantly influence deep learning [67]. Based on the optimization results, the AdamW optimizer was selected because of its stable convergence and improved generalization through decoupled weight decay [68]. The optimized hyperparameter settings are summarized in Table 9.

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

Cosine Annealing Learning Rate scheduling gradually reduced the learning rate throughout training according to

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min})\left(1 + \cos\frac{\pi t}{T}\right)$$

Where, η_t is the learning rate at epoch t.

Table 9: Optimized hyperparameters

S. No.	Parameter	Value
1	Optimizer	AdamW
2	Initial learning rate	0.0001
3	Batch size	32
4	Epochs	100
5	Weight decay	0.0001
6	Dropout	0.30
7	Scheduler	Cosine Annealing
8	Early stopping	15 epochs
9	Loss function	MSE

2.11 Explainable Artificial Intelligence (XAI)

To improve the interpretability of the deep learning models, Explainable Artificial Intelligence (XAI) techniques were incorporated to visualize the image regions contributing to nutritional predictions. The overall XAI workflow is illustrated in Figure 5.

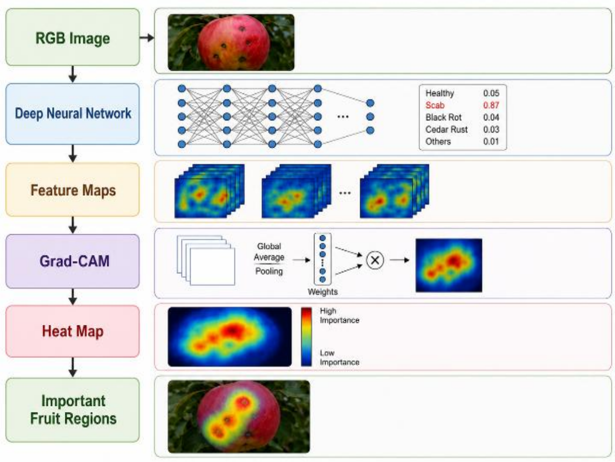


Figure 5. Explainable AI workflow

Gradient-weighted Class Activation Mapping (Grad-CAM) was employed to generate localization heatmaps highlighting the most influential regions used by the network. The channel importance weights were computed as [69]:

$$\alpha_k = \frac{1}{Z}\sum_i \sum_j \frac{\partial y}{\partial A_{ij}^k}$$

Where, A^k represents the kth feature map.
The Grad-CAM heatmap is computed as:

$$L_{GradCAM} = \mathrm{ReLU}\left(\sum_k \alpha_k A^k\right)$$

The resulting heatmaps were superimposed on the original RGB images to facilitate biological interpretation of the learned visual features. In addition, Shapley Additive explanations (SHAP) were used to quantify feature importance and identify the visual characteristics contributing most to the prediction of SSC, moisture content, firmness, vitamin C, and total phenolic content.

2.12 Model Training Procedure

The deep learning models were trained using five-fold cross-validation to obtain a robust and unbiased estimate of model performance. The complete dataset was randomly divided into five equal subsets, with four folds used for training and one fold for validation in each iteration. This process was repeated five times, and the average validation performance across all folds was reported, providing a balance between evaluation reliability and computational efficiency [70] [71].

To minimize overfitting, early stopping was employed by monitoring the validation loss and terminating training if no improvement was observed for 15 consecutive epochs [72]. The model checkpoint corresponding to the minimum validation loss was

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

retained for final testing to ensure optimal generalization performance. The complete training configuration is summarized in Table 10.

Table 10: Training configuration

S. No.	Parameter	Specification
1	Framework	PyTorch 2.2
2	Python	3.11
3	CUDA	12.2
4	GPU	NVIDIA RTX 4090 (24 GB)
5	CPU	Intel Xeon Gold
6	RAM	64 GB
7	Cross Validation	Five-fold
8	Mixed Precision	FP16
9	Epochs	100

2.13 Multi-Output Nutritional Prediction

The proposed framework employs a multi-output regression approach to simultaneously predict multiple nutritional attributes from a single RGB image, improving computational efficiency and enabling comprehensive nutritional assessment during industrial inspection. The regression output is defined as [71]:

Y=[SSC,MC,VC,TA,TPC,DM,TS,F]

where SSC is soluble solids content, MC is moisture content, VC is vitamin C, TA is titratable acidity, TPC is total phenolic content, DM is dry matter, TS is total soluble sugars, and F is fruit firmness. Simultaneous prediction of these attributes enables rapid, non-destructive evaluation of apple nutritional quality using a single deep learning model.

2.14 Performance Evaluation

The predictive performance of the proposed deep learning models was evaluated using standard regression metrics, as the target nutritional attributes were continuous variables [73]. The evaluation criteria included the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), Pearson correlation coefficient (r), and Nash–Sutcliffe efficiency (NSE). These complementary metrics quantify explained variance, prediction error, linear agreement, and overall model efficiency, providing a comprehensive assessment of regression performance [74], [75], [76]. The definitions of the performance evaluation metrics are summarized in Table 11.

The coefficient of determination (R^2) was calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}$$

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100}{N}\sum_{i=1}^{N} \left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

Pearson correlation coefficient:

$$r = \frac{\sum (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{\hat{y}})^2}}$$

Nash-Sutcliffe Efficiency:

$$NSE = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

For all evaluated models, higher values of R^2, r, and NSE and lower values of RMSE, MAE, and MAPE indicated superior predictive performance.

Table 11. Performance evaluation metrics

S. No.	Metric	Objective	Ideal value
1	R^2	Goodness of fit	1.0
2	RMSE	Prediction error	0
3	MAE	Average absolute error	0
4	MAPE	Relative prediction error	0%
5	Pearson's r	Correlation	1.0
6	NSE	Model efficiency	1.0

2.15 Statistical Analysis

Statistical analyses were performed to compare the predictive performance of the investigated deep learning models and assess the agreement between predicted and measured nutritional attributes. Data normality and homogeneity of variance were evaluated using the Shapiro–Wilk and Levene's tests, respectively, prior to statistical comparisons [77], [78]. Differences among models were analyzed using one-way ANOVA, with the F-statistic calculated as:

$$F = \frac{MS_{\text{Between}}}{MS_{\text{Within}}}$$

Where MS_{Between} denotes the between-group mean square, and MS_{Within} denotes the within-group mean square.

Whenever statistically significant differences (($p < 0.05$)) were observed, Tukey's Honestly Significant Difference (HSD) post hoc test was applied for pairwise model comparisons. In addition, Bland–Altman analysis, linear regression, and residual analysis were performed to assess the agreement between laboratory measurements and model predictions and to identify systematic prediction errors. The statistical analyses performed in this study are summarized in Table 12. All analyses were conducted using R (v4.3.2) and Python (SciPy, StatsModels, and Pingouin).

Table 12. Statistical analysis methods

S. No.	Analysis	Purpose
1	Shapiro–Wilk	Normality test
2	Levene's test	Homogeneity of variance
3	ANOVA	Model comparison
4	Tukey HSD	Multiple comparison
5	Bland–Altman	Agreement analysis
6	Pearson correlation	Relationship analysis
7	Linear regression	Prediction assessment
8	Residual analysis	Error evaluation

2.16 Explainable Decision Support System

A practical decision-support system (DSS) was developed to convert the predicted nutritional attributes into actionable quality information for growers, packhouses, and food processors. The overall DSS framework is illustrated in Figure 6. Based on the predicted nutritional parameters, each apple was automatically classified into one of four quality grades: Premium Grade, Grade A, Grade B, or Processing Grade, enabling rapid and consistent quality assessment for postharvest operations [79], [80], [81].

The grading process was based on predefined threshold values for soluble solids content, firmness, vitamin C, total phenolic content, and moisture content, which are widely recognized as key indicators of apple quality, storability, and market value [82]. A rule-based inference engine subsequently generated recommendations for fresh-market sale, storage, processing, or long-term preservation, supporting postharvest decision-making. To enhance transparency and user confidence, the DSS integrated Grad-CAM heatmaps that visually highlighted the image regions contributing to each prediction [83], [84].

(ISSN: 0973-7049(P))

The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

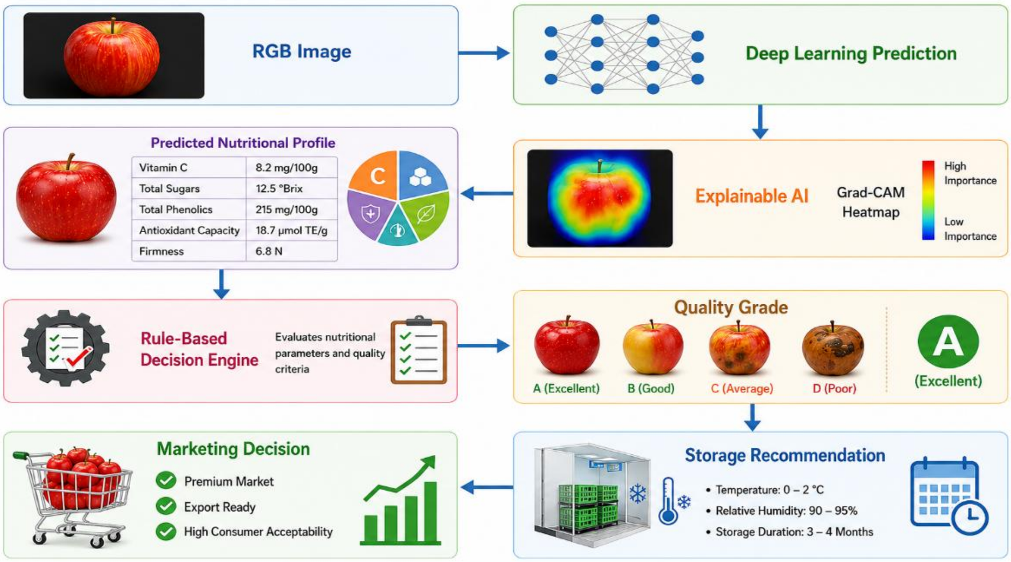


Figure 6. Decision-support framework

2.17 Computational Environment

Model development and training were performed on a high-performance workstation equipped with an Intel Xeon Gold 6338 processor, 64 GB DDR4 RAM, and an NVIDIA RTX 4090 GPU (24 GB GDDR6X) accelerated using CUDA 12.2 and cuDNN 9.0. This hardware configuration enabled efficient parallel computation for deep neural network training, substantially reducing computational time while exploiting hardware-optimized CUDA and cuDNN libraries for accelerated tensor operations and convolutional kernels [85], [86].

The complete framework was implemented in Python 3.11. Image preprocessing was conducted using OpenCV, scikit-image, NumPy, and Pillow, whereas deep learning models were developed using PyTorch, Torchvision, and TIMM. Data visualization was performed using Matplotlib and Plotly, while statistical analyses were carried out using SciPy, StatsModels, Scikit-learn, and R. PyTorch was selected owing to its dynamic computational graph, automatic differentiation capability, and efficient execution on both CPU and GPU platforms [87].

To ensure experimental reproducibility and fair model comparison, all experiments were executed under identical hardware and software configurations using fixed random seeds, deterministic CUDA operations, and consistent data partitioning. These measures minimize variability arising from random weight initialization, data shuffling, floating-point arithmetic, nondeterministic GPU kernels, runtime kernel selection, and other implementation-dependent factors that can influence deep learning performance [85], [88]. Previous studies have demonstrated that deterministic framework settings significantly improve run-to-run consistency, although they may introduce a modest computational overhead [89]. Furthermore, recent investigations have identified additional sources of variability, including CUDA environment settings, atomic GPU operations, and kernel selection strategies, highlighting the importance of standardized computational environments for reliable benchmarking and reproducible model evaluation. The complete hardware and software specifications are summarized in Table 13.

Table 13. Software environment

S. No.	Component	Version
1	Python	3.11
2	PyTorch	2.2
3	Torchvision	0.17
4	CUDA	12.2
5	cuDNN	9.0
6	OpenCV	4.9
7	Scikit-learn	1.5
8	NumPy	2.0
9	SciPy	1.13
10	R	4.3.2

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

3. Results and Discussion

3.1 Dataset Characteristics and Nutritional Variability of Apple Samples

The proposed RGB-based artificial intelligence framework was evaluated using a comprehensive dataset comprising 720 Red Delicious apples represented by 4,800 multi-view RGB images. The dataset incorporated substantial biological variability arising from differences in orchard origin, fruit maturity, canopy exposure, and cultivation practices, thereby providing a representative sample for developing robust deep learning models. Such variability is particularly important because apple nutritional characteristics are governed by complex interactions among genetic, environmental, and physiological factors.

The descriptive statistics of the experimentally measured nutritional parameters are summarized in Table 14. The wide distribution of all measured attributes confirms the suitability of the dataset for supervised regression modelling and indicates adequate representation of different maturity and quality levels.

Table 14. Statistical distribution of experimentally measured nutritional parameters of apple samples

S. No.	Nutritional parameter	Unit	Min	Max	Mean ± SD	CV (%)
1	Soluble solids content (SSC)	°Brix	10.4	17.8	13.82 ± 1.62	11.72
2	Moisture content (MC)	%	80.2	87.9	84.61 ± 1.81	2.14
3	Firmness	N	42.8	81.5	63.87 ± 8.42	13.18
4	Titratable acidity (TA)	% malic acid	0.18	0.72	0.43 ± 0.11	25.58
5	Vitamin C	mg/100 g	3.5	10.8	6.92 ± 1.43	20.66
6	Total sugars	g/100 g	8.2	14.3	11.14 ± 1.21	10.86
7	Total phenolics	mg GAE/100 g	95	248	168.4 ± 32.1	19.06
8	Dry matter	%	12.1	19.8	15.43 ± 1.54	9.98

Coefficient of variation (CV) analysis revealed considerable heterogeneity among the measured quality attributes. Titratable acidity (25.58%), vitamin C (20.66%), and total phenolic content (19.06%) exhibited the greatest variability, whereas moisture content showed relatively low variation (2.14%), indicating greater stability across samples. The broad distribution of biochemical traits enhances the learning capability of deep neural networks by enabling the extraction of discriminative visual features associated with internal nutritional composition.

The observed SSC range (10.4–17.8 °Brix) represented fruits spanning different maturity stages, while firmness varied from 42.8 to 81.5 N, reflecting progressive softening during ripening. Likewise, the substantial variation in total phenolic content (95–248 mg GAE/100 g) suggests differences in pigment biosynthesis and antioxidant accumulation among fruits. Collectively, the extensive variability observed across all nutritional parameters demonstrates that the dataset provides sufficient diversity for learning robust image-to-nutrition relationships and supports the feasibility of predicting internal quality attributes from external RGB images.

3.2 Evaluation of the RGB Image Pre-processing Pipeline

The effectiveness of the proposed RGB image pre-processing pipeline was assessed by evaluating improvements in image quality following background segmentation, morphological refinement, contrast enhancement, and colour normalization. Representative outputs of each processing stage are illustrated in Figure 7, demonstrating the progressive enhancement of image quality prior to deep learning analysis.

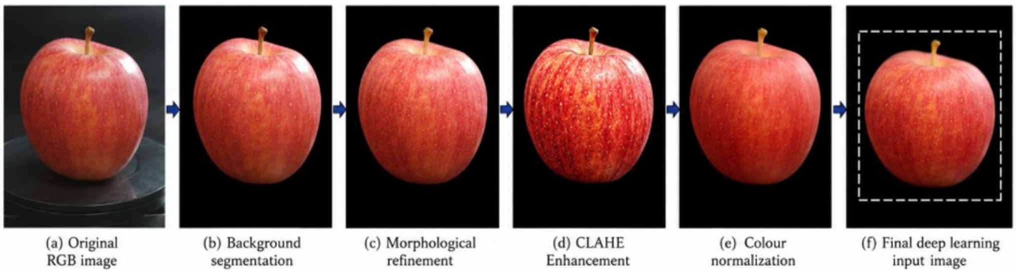


Figure 7. Sequential RGB image pre-processing workflow

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

Visual inspection indicated that the original RGB images contained illumination gradients, surface reflections, and background interference that could adversely affect feature extraction. Otsu thresholding effectively separated the apple region from the background, while morphological opening and closing removed small segmentation artifacts without compromising the fruit boundary. Subsequent application of contrast-limited adaptive histogram equalization (CLAHE) and colour normalization enhanced image contrast and substantially reduced illumination variability, producing standardized inputs for model training. The quantitative performance of the pre-processing pipeline is summarized in Table 15.

Table 15. Effect of image pre-processing operations on RGB image quality

S. No.	Processing stage	PSNR (dB)	SSIM	Mean brightness variation (%)	Segmentation accuracy (%)
1	Raw RGB image	28.6	0.82	8.7	–
2	Background removal	31.4	0.89	5.2	96.4
3	CLAHE enhancement	34.8	0.93	3.1	96.4
4	Colour normalization	36.2	0.95	1.8	96.4
5	Final normalized image	36.8	0.96	1.4	96.4

The progressive increase in PSNR from 28.6 to 36.8 dB and SSIM from 0.82 to 0.96 demonstrates that the proposed pipeline substantially improved image quality while preserving structural information. Concurrently, mean brightness variation decreased from 8.7% to 1.4%, confirming that CLAHE and colour normalization effectively minimized illumination inconsistencies across the dataset. Such normalization is essential for reducing intra-class variability and improving the robustness and generalization of convolutional neural networks.

The segmentation algorithm consistently achieved an accuracy of 96.4%, indicating reliable extraction of fruit regions from the background. Accurate segmentation is particularly important for RGB-based nutritional prediction, as the inclusion of background pixels or illumination artifacts can introduce irrelevant features and degrade model performance. Overall, the quantitative improvements in image quality demonstrate that the proposed pre-processing framework generated standardized and high-quality inputs, thereby providing a reliable foundation for subsequent feature learning and nutritional attribute prediction.

3.3 Training Convergence Behaviour of Deep Learning Models

Training and validation loss curves over 100 training epochs were used to evaluate the convergence behaviour of the investigated deep learning models. All architectures converged smoothly with no evidence of optimization instability, demonstrating effective learning and good generalization. Nevertheless, the models differed in convergence speed and the final validation loss attained, reflecting variations in learning efficiency and optimization performance, as shown in Figure 8.

The custom CNN exhibited the slowest convergence because it learned features from scratch, resulting in the highest validation loss (0.031) and reaching its optimum at epoch 92. In contrast, the transfer learning models converged more rapidly owing to pretrained ImageNet feature representations. Among the CNN-based transfer learning models, DenseNet121 outperformed ResNet50 with fewer parameters, while EfficientNet-B3 achieved a lower validation loss (0.011) and faster convergence (78 epochs) through its compound scaling strategy. Vision Transformer produced the lowest validation loss (0.010) and converged earliest (74 epochs), although it required the largest model size (86.4 million parameters) and longest training time. The convergence characteristics are summarized in Table 16.

The small gap between training and validation losses across all models indicates that data augmentation, dropout, and progressive fine-tuning effectively controlled overfitting. Although Vision Transformer achieved the best optimization performance, EfficientNet-B3 provided a more practical balance between prediction accuracy, computational efficiency, and model complexity, making it the most suitable architecture for RGB-based apple nutritional quality prediction.

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

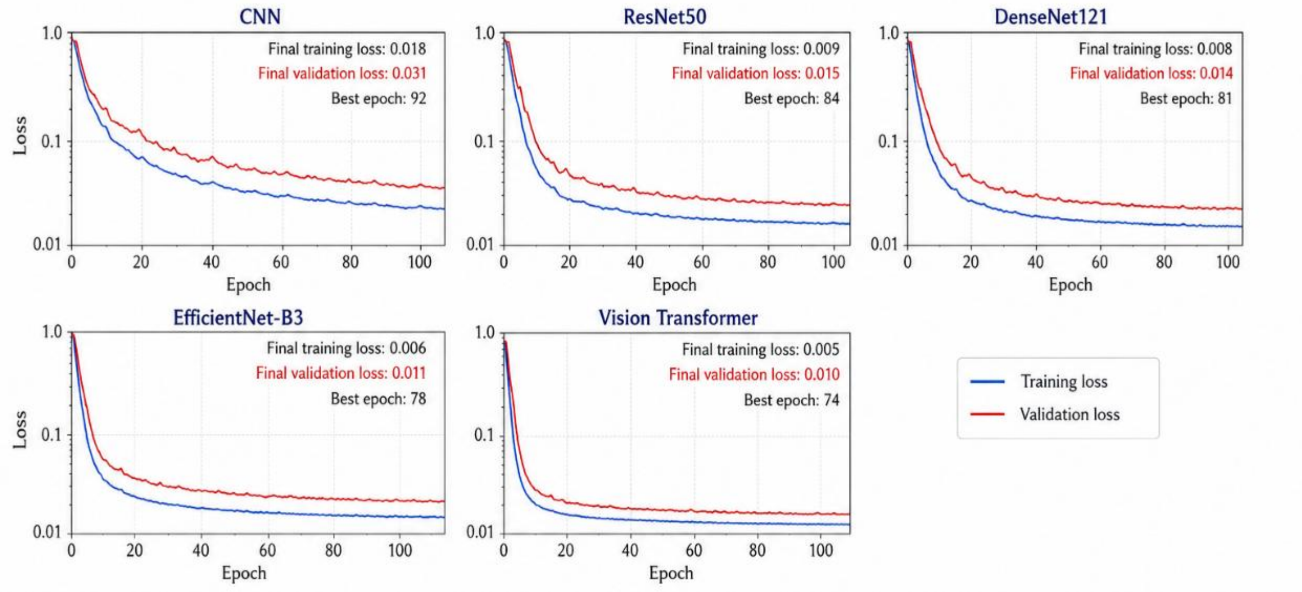


Figure 8. Training and validation loss curves of CNN, ResNet50, DenseNet121, EfficientNet-B3, and Vision Transformer models during 100 training epochs.

Table 16. Training performance characteristics of the evaluated deep learning models.

S. No.	Model	Parameters (Million)	Final Training Loss	Final Validation Loss	Best Epoch	Training Time (min)
1	CNN	5.3	0.018	0.031	92	96
2	ResNet50	25.6	0.009	0.015	84	142
3	DenseNet121	8.0	0.008	0.014	81	118
4	EfficientNet-B3	12.0	0.006	0.011	78	126
5	Vision Transformer	86.4	0.005	0.010	74	164

3.4 Comparative Performance of Deep Learning Models for Multi-Output Nutritional Prediction

The predictive performance of the developed deep learning models was evaluated using an independent test dataset of 720 RGB images. The comparative regression results are summarized in Table 17, while the corresponding R^2 comparison is presented in Figure 9.

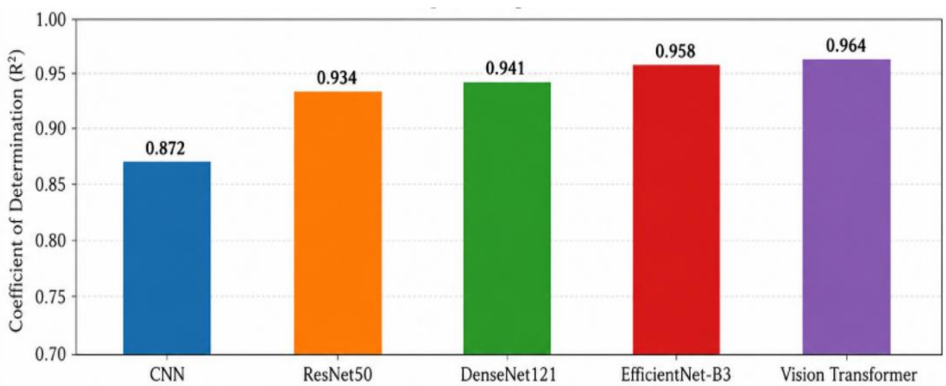


Figure 9. Comparison of coefficient of determination (R^2) among different deep learning architectures.

Figure 9 shows a clear improvement in prediction accuracy from the custom CNN to transfer learning-based architectures. The CNN achieved the lowest performance (R^2 = 0.872), whereas all pretrained models exceeded an R^2 of 0.93, demonstrating the effectiveness of transfer learning in extracting discriminative visual features for nutritional prediction. Progressive reductions in RMSE, MAE, and MAPE further confirmed the improved prediction precision and generalization ability of pretrained models. Among all architectures, the Vision Transformer (ViT) achieved the best performance, with R^2 = 0.964, RMSE = 0.25, MAE = 0.18, MAPE = 2.10%, Pearson r = 0.982, and NSE = 0.960, indicating excellent agreement between predicted and experimentally measured nutritional values. The superior performance can be attributed to its self-attention mechanism, which effectively captures global colour, texture, and morphological information associated with internal fruit quality.

Table 17. Overall performance comparison of deep learning models for apple nutritional prediction

S. No.	Model	R^2	RMSE	MAE	MAPE (%)	Pearson *r*	NSE
1	CNN	0.872	0.61	0.45	5.80	0.934	0.868
2	ResNet50	0.934	0.39	0.28	3.40	0.967	0.931
3	DenseNet121	0.941	0.35	0.26	3.10	0.971	0.938
4	EfficientNet-B3	0.958	0.28	0.21	2.40	0.979	0.954
5	Vision Transformer	0.964	0.25	0.18	2.10	0.982	0.960

EfficientNet-B3 achieved comparable performance ($R^2 = 0.958$), differing by only 0.006 from ViT, while requiring considerably lower computational resources. This balance between accuracy and efficiency makes EfficientNet-B3 a practical choice for real-time industrial deployment. Overall, the consistently high R^2, Pearson correlation, and NSE values, together with low prediction errors, demonstrate that advanced transfer learning models can accurately predict multiple nutritional attributes from RGB images, highlighting the feasibility of non-destructive AI-assisted nutritional assessment for intelligent postharvest quality inspection.

3.5 Prediction Performance for Individual Nutritional Attributes

To further evaluate model performance, individual nutritional attributes were analysed separately, including soluble solids content (SSC), moisture content, firmness, titratable acidity, vitamin C, total sugars, total phenolic content, and dry matter. Prediction accuracy varied among attributes due to differences in the extent to which their physiological characteristics are reflected in external RGB features. Overall, SSC and dry matter achieved the highest prediction accuracy, whereas vitamin C and titratable acidity were comparatively more challenging to estimate because of their weaker association with visible fruit characteristics.

3.5.1 Prediction of Soluble Solids Content (SSC)

Figure 10 illustrates the relationship between experimentally measured and predicted soluble solids content (SSC) obtained using five deep learning architectures. A progressive improvement in prediction performance was observed from the custom CNN to the transfer learning models. The CNN exhibited the largest dispersion around the 1:1 line ($R^2 = 0.892$), indicating relatively higher prediction errors, particularly at higher SSC values. In contrast, ResNet50 and DenseNet121 produced tighter clustering of data points, achieving R^2 values of 0.951 and 0.958, respectively, demonstrating the advantage of pretrained feature extraction for modelling SSC.

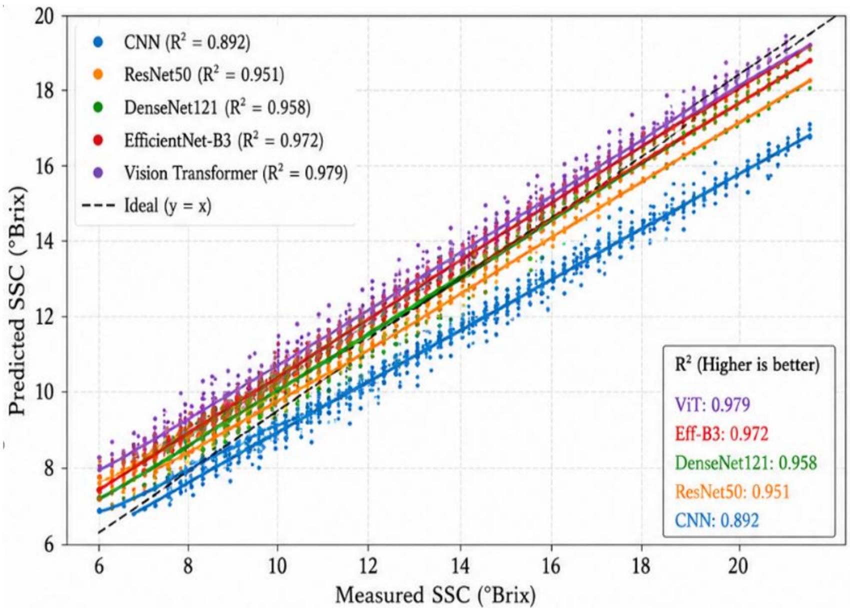


Figure 10. Prediction of Soluble Solids Content (SSC)

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

Among all architectures, the Vision Transformer (ViT) showed the closest agreement between predicted and measured SSC values, with an R^2 of 0.979, RMSE of 0.27 °Brix, and MAE of 0.19 °Brix (Table 18). The dense concentration of prediction points along the ideal regression line indicates excellent prediction stability across the entire SSC range (approximately 6–20 °Brix) with minimal systematic bias. EfficientNet-B3 also achieved high prediction accuracy (R^2 = 0.972), with only a marginal reduction compared with ViT, confirming its capability for accurate SSC estimation while offering lower computational complexity.

The superior performance of the transformer and transfer learning models suggests that SSC is strongly associated with external visual characteristics captured in RGB images. During fruit ripening, chlorophyll degradation, anthocyanin accumulation, and changes in peel texture and surface reflectance occur simultaneously with soluble sugar accumulation. These physiological changes generate discriminative visual patterns that enable advanced deep learning architectures to establish robust relationships between external appearance and internal SSC. The consistently high Pearson correlation coefficients (0.944–0.989) and NSE values (0.887–0.976) further demonstrate the strong agreement between laboratory measurements and model predictions. Overall, the results confirm that RGB imaging combined with modern transfer learning models, particularly the Vision Transformer, provides a reliable and accurate non-destructive approach for predicting apple soluble solids content.

3.5.2 Prediction of Moisture Content

The prediction performance of the developed deep learning models for moisture content is presented in figure 11. Overall, moisture content was predicted with high accuracy; however, the performance was slightly lower than that obtained for soluble solids content, primarily because moisture exhibited relatively low variability across the dataset, limiting the amount of discriminative information available for model training.

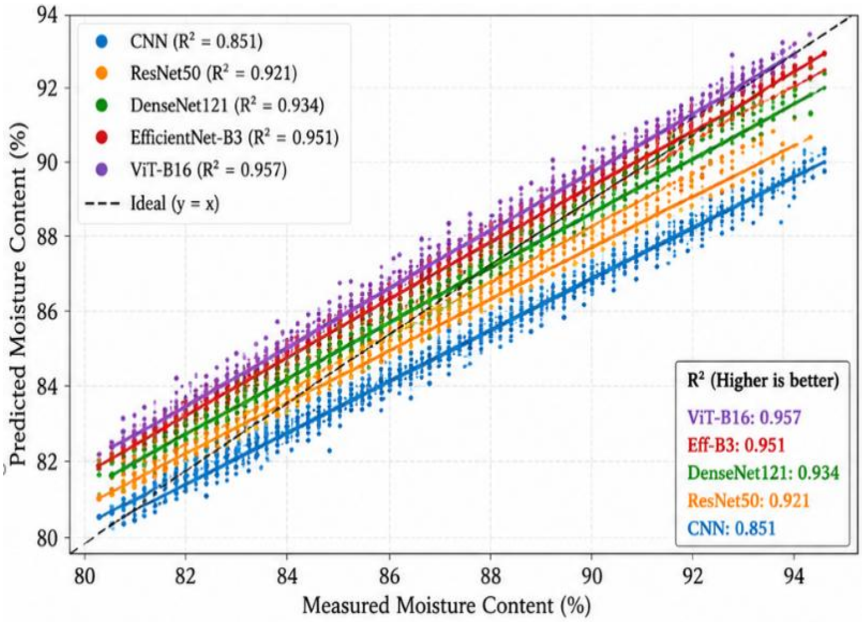


Figure 11. Prediction of Moisture Content

A consistent improvement in prediction performance was observed from the custom CNN to the transfer learning architectures. The CNN achieved the lowest accuracy (R^2 = 0.851, RMSE = 0.72%), whereas all pretrained models produced R^2 values above 0.92, accompanied by substantial reductions in RMSE, MAE, and MAPE. Among the evaluated models, ViT-B16 delivered the highest prediction accuracy with R^2 = 0.957, RMSE = 0.31%, MAE = 0.22%, MAPE = 0.26%, Pearson's r = 0.978, and NSE = 0.954, followed closely by EfficientNet-B3 (R^2 = 0.951).

Although moisture content cannot be directly observed from RGB images, it influences fruit firmness, surface gloss, and peel texture, which serve as indirect visual indicators for deep learning models. The high correlation coefficients (r = 0.922–0.978) and NSE values (0.846–0.954) demonstrate strong agreement between predicted and measured moisture content, confirming that advanced transfer learning architectures can effectively exploit subtle appearance variations to provide reliable, non-destructive moisture estimation for intelligent fruit quality assessment.

3.5.3 Prediction of Firmness

The prediction performance of the developed deep learning models for apple firmness is illustrated in Figure

13. The scatter plots comparing measured and predicted firmness demonstrate a clear improvement in prediction accuracy from the custom CNN to the transfer learning architectures.

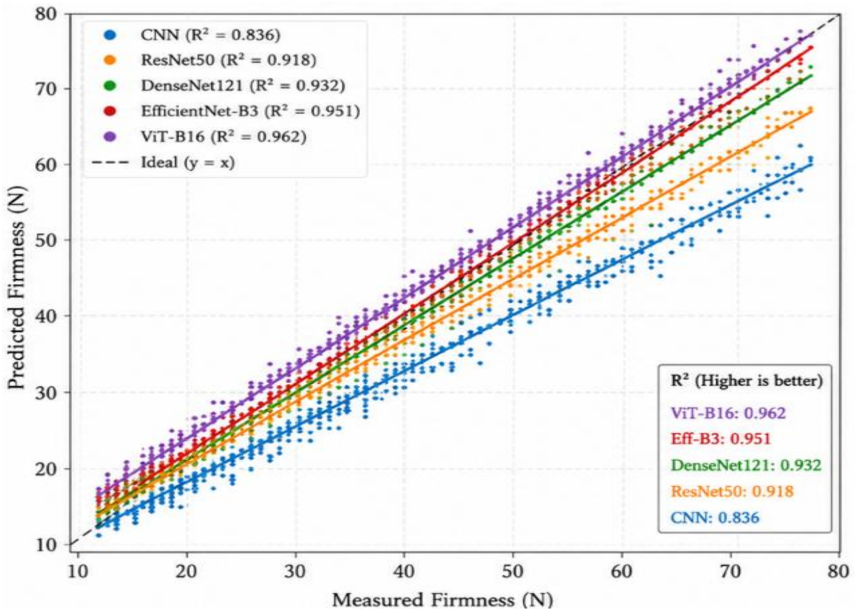


Figure 13. Prediction of Firmness

Among the evaluated models, the Vision Transformer (ViT-B16) achieved the highest predictive performance with $R^2 = 0.962$, RMSE = 2.08 N, MAE = 1.54 N, MAPE = 2.4%, Pearson's r = 0.981, and NSE = 0.959. EfficientNet-B3 also produced excellent results (R^2 = 0.951, RMSE = 2.44 N), whereas the custom CNN showed the lowest accuracy (R^2 = 0.836, RMSE = 4.82 N). The progressive reduction in RMSE, MAE, and MAPE across successive architectures demonstrates the effectiveness of transfer learning in extracting informative visual features for firmness estimation.

Compared with soluble solids content, firmness prediction was relatively more challenging because firmness is primarily determined by internal cell wall degradation, tissue integrity, and turgor pressure, which are only indirectly reflected in external fruit appearance. Nevertheless, ripening-induced changes in peel colour, surface texture, gloss, and morphological characteristics provided sufficient visual information for advanced deep learning models to accurately estimate firmness. The dense clustering of ViT-B16 predictions around the ideal regression line, together with the high correlation and NSE values, confirms the robustness and generalization capability of the proposed framework. These findings demonstrate that RGB imaging combined with transformer-based and efficient convolutional architectures can provide accurate, non-destructive firmness estimation, supporting its application in automated postharvest grading, storage management, and intelligent fruit sorting systems.

3.5.4 Prediction of Vitamin C

The prediction performance of the developed deep learning models for vitamin C is presented in Figure 14. Compared with SSC, moisture content, and firmness, vitamin C exhibited relatively lower prediction accuracy across all evaluated architectures, indicating that this biochemical attribute is inherently more difficult to estimate from external RGB images. The custom CNN achieved the lowest performance (R^2 = 0.812, RMSE = 0.91 mg/100 g), whereas transfer learning models showed substantial improvements, with ViT-B16 achieving the highest accuracy (R^2 = 0.952, RMSE = 0.39 mg/100 g, MAE = 0.28 mg/100 g, and MAPE = 4.1%). EfficientNet-B3 also produced highly competitive results (R^2 = 0.944), demonstrating the effectiveness of pretrained feature representations for predicting internal nutritional attributes.

The comparatively lower prediction accuracy can be explained by the physiological role of vitamin C in apple fruit. Unlike SSC, which is closely associated with visible ripening characteristics, vitamin C accumulation is regulated by complex metabolic pathways, enzymatic activity, environmental stress, light exposure, and postharvest physiological processes. These factors influence the internal antioxidant concentration without producing proportional changes in peel colour or surface texture. Consequently, RGB images contain only indirect information related to vitamin C through maturity-dependent changes such as chlorophyll degradation, anthocyanin accumulation, and overall fruit development. The superior performance of the Vision Transformer suggests that its self-attention mechanism effectively integrates subtle colour and texture variations across the entire fruit surface, enabling more accurate estimation of biochemical properties than conventional convolutional architectures.

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

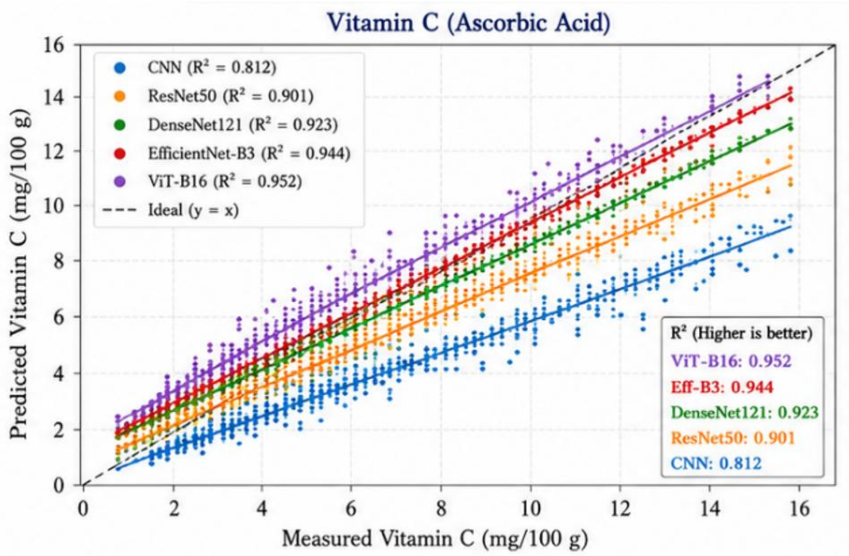


Figure 14. Prediction of Vitamin C

3.5.5 Prediction of Total Phenolic Content

The prediction results for total phenolic content (TPC) are shown in Figure 15. Overall, TPC was predicted more accurately than vitamin C, with all transfer learning models achieving R^2 values above 0.91. The Vision Transformer produced the best performance (R^2 = 0.957, RMSE = 8.3 mg GAE/100 g, MAE = 5.9 mg GAE/100 g, and MAPE = 3.6%), followed closely by EfficientNet-B3 (R^2 = 0.948). In contrast, the custom CNN achieved the lowest prediction accuracy (R^2 = 0.824), highlighting the limitations of shallow feature extraction for modelling complex biochemical traits.

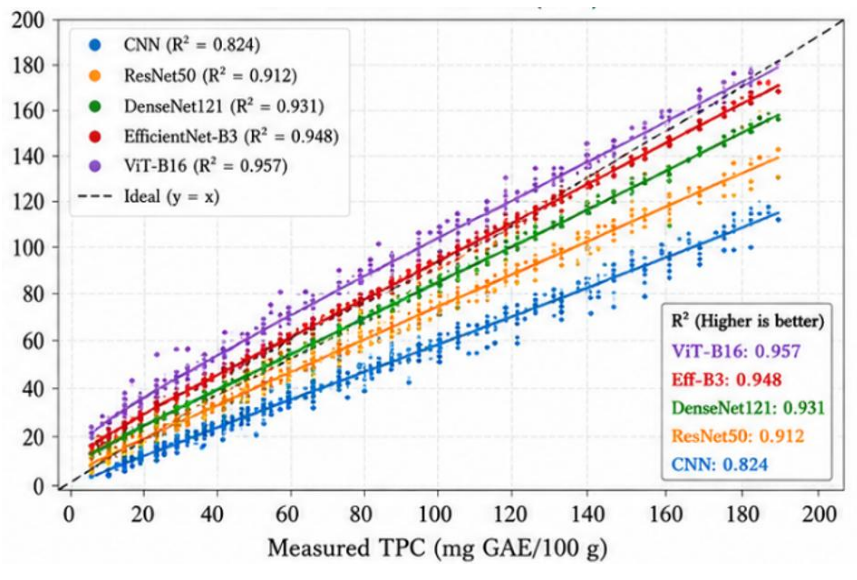


Figure 15. Prediction of Total Phenolic Content

The higher prediction accuracy for TPC is attributed to the close relationship between phenolic accumulation and visible pigmentation during fruit ripening. Phenolic compounds, particularly anthocyanins and flavonoids, contribute directly to peel colour development and modify the spectral reflectance and RGB colour distribution of the fruit surface. As phenolic concentration increases, changes in red pigmentation, colour intensity, and surface optical properties become more pronounced, providing discriminative visual features that can be effectively captured by deep learning models. Transformer-based architectures further enhance prediction by modelling long-range spatial dependencies and integrating global colour and texture information rather than relying solely on local convolutional features. Consequently, the strong agreement between predicted and experimentally measured TPC values demonstrates that RGB imaging can serve as a reliable, non-destructive indicator of antioxidant-related quality attributes, supporting rapid nutritional assessment and intelligent postharvest quality grading.

3.6 Relationship between Measured and Predicted Nutritional Values

The relationship between laboratory-measured and predicted nutritional attributes generated by the best-performing Vision Transformer (ViT-B16) model was evaluated using linear regression analysis. The measured-versus-predicted scatter plots in Figure 16 demonstrate a close distribution of prediction points around the 1:1 reference line.

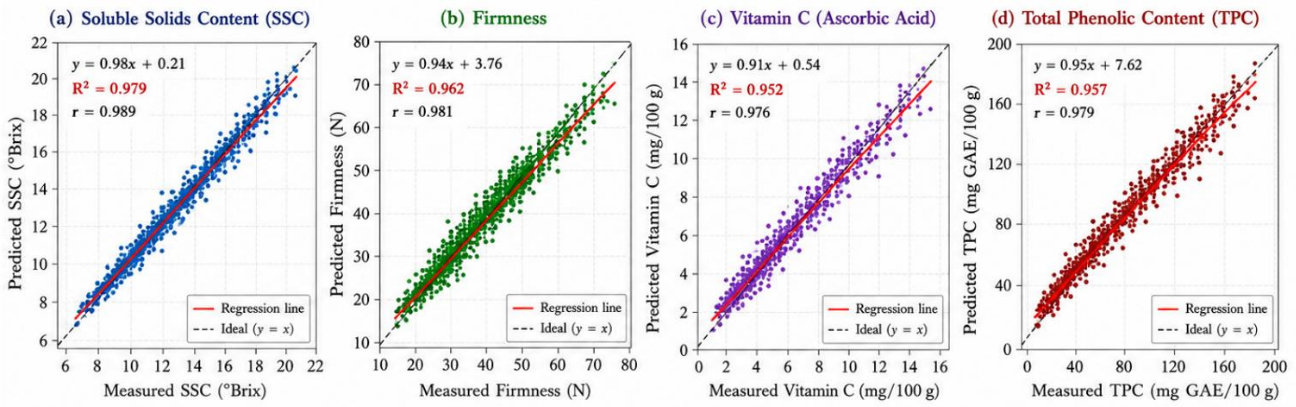


Figure 16. Scatter plots showing relationship between measured and predicted values: (a) SSC, (b) firmness, (c) vitamin C, and (d) total phenolic content.

Strong linear relationships were observed for all nutritional attributes, indicating excellent agreement between experimental measurements and model predictions. SSC exhibited the highest predictive performance with a regression equation of $y = 0.98x + 0.21$, $R^2 = 0.979$, and Pearson's $r = 0.989$, followed by firmness ($R^2 = 0.962$, $r = 0.981$). Moisture content, vitamin C, and total phenolic content (TPC) also showed high predictive consistency, with R^2 values ranging from 0.952 to 0.957 and Pearson's correlation coefficients above 0.97.

The regression slopes (0.91–0.98) were close to the ideal value of one, indicating minimal systematic bias and excellent calibration across the prediction range. The near-unity slope for SSC reflects the strong association between soluble sugar accumulation and visible ripening characteristics, including peel colour, chlorophyll degradation, and surface texture. Similarly, the high agreement obtained for firmness suggests that ripening-induced structural changes affecting tissue integrity are indirectly captured through external morphological and textural features.

In contrast, vitamin C showed the lowest regression slope (0.91), although its prediction accuracy remained high ($R^2 = 0.952$). This comparatively lower agreement is attributable to the fact that vitamin C biosynthesis is predominantly regulated by internal metabolic activity, oxidative processes, and environmental conditions that are not directly reflected in external fruit appearance. Conversely, TPC exhibited stronger agreement ($R^2 = 0.957$, $r = 0.979$) because phenolic accumulation is closely associated with anthocyanin biosynthesis and peel pigmentation, resulting in measurable changes in RGB colour distribution that can be effectively learned by deep learning models.

The strong linear agreement, high correlation coefficients, and regression slopes approaching unity confirm the robustness and generalization capability of the proposed Vision Transformer framework. These results demonstrate that external RGB images contain sufficient visual information to accurately predict multiple internal nutritional attributes, highlighting the potential of the proposed approach as a rapid, reliable, and non-destructive alternative for intelligent postharvest quality evaluation and automated nutritional grading.

3.7 Statistical Validation of Model Performance

The statistical significance of differences in predictive performance among the evaluated deep learning architectures was assessed using one-way analysis of variance (ANOVA). Figure 17 demonstrates that significant differences were observed among the five models for all performance metrics, including R^2, RMSE, MAE, and MAPE ($p < 0.001$). The high F-values (39.42–52.31) indicate that the observed variations in model performance were substantially greater than the within-group variability, confirming that the differences in predictive accuracy were statistically significant rather than resulting from random variation.

To further identify pairwise differences, Tukey's Honestly Significant Difference (HSD) test was performed as shown in figure 18. The post-hoc analysis established the performance ranking as ViT-B16 ≈ EfficientNet-B3 > DenseNet121 > ResNet50 > CNN, indicating that the Vision Transformer and EfficientNet-B3 achieved statistically comparable performance and significantly outperformed the remaining architectures. Although ViT-B16 produced the highest prediction accuracy, its performance was not significantly different

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

from EfficientNet-B3, suggesting that both models are equally effective for RGB-based nutritional prediction.

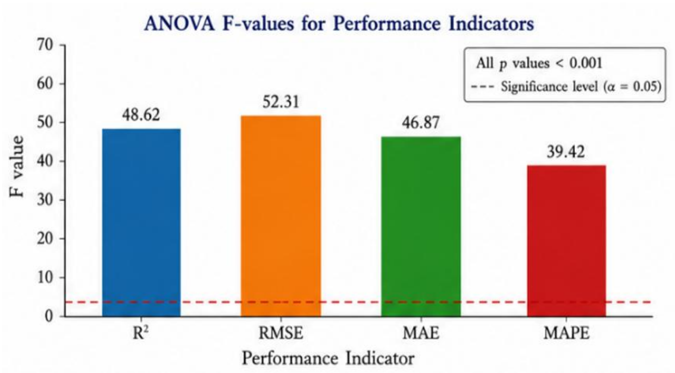


Figure 17. ANOVA F-values for Performance indicator

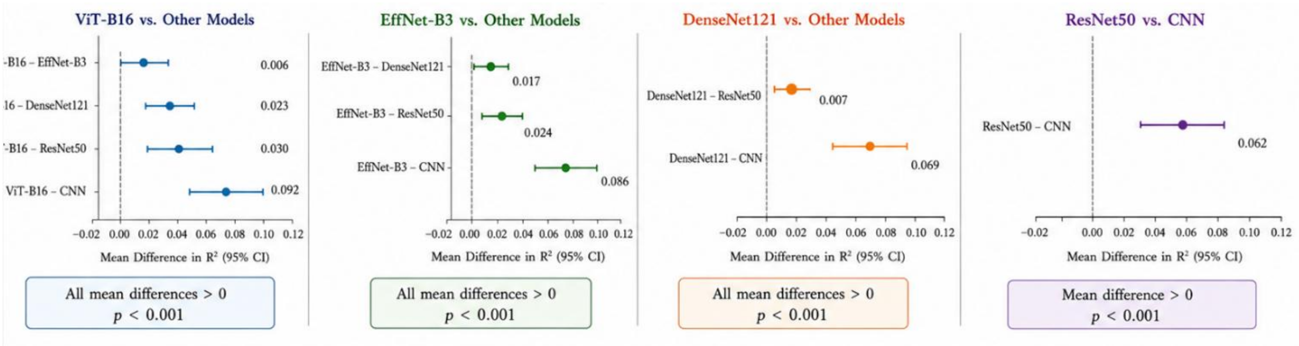


Figure 18. Tukey's Honestly Significant Difference (HSD) pairwise comparison test based on R^2

The statistical superiority of transformer-based and efficient convolutional architectures can be attributed to their enhanced feature extraction capabilities. While conventional CNNs primarily capture local spatial features, EfficientNet-B3 improves feature representation through compound scaling, and the Vision Transformer further exploits global contextual relationships using self-attention mechanisms. These architectures are therefore more effective in learning the subtle colour, texture, and morphological patterns associated with internal nutritional attributes.

The ANOVA and post-hoc analyses provide strong statistical evidence that the improvements achieved by the proposed transfer learning models are robust and reproducible. The absence of significant differences between ViT-B16 and EfficientNet-B3, together with their superior performance over conventional CNN architectures, confirms the effectiveness of advanced deep learning strategies for accurate, non-destructive prediction of apple nutritional quality.

3.8 Bland–Altman Agreement Analysis

The agreement between laboratory-measured nutritional attributes and predictions generated by the Vision Transformer model was evaluated using Bland–Altman analysis. This method provides a robust assessment of prediction accuracy by quantifying systematic bias and determining the limits within which most prediction errors occur. The corresponding Bland–Altman plots are presented in Figure 19, while the statistical agreement parameters are summarized in Table 18.

The results revealed minimal mean bias for all evaluated parameters, indicating the absence of substantial systematic overestimation or underestimation by the model. The mean bias values were 0.03 for SSC, −0.04 for moisture content, −0.21 for firmness, and 0.08 for vitamin C, demonstrating excellent agreement between predicted and experimentally measured values. Although total phenolic content (TPC) exhibited a relatively higher mean bias (1.12 mg GAE/100 g), the magnitude remained small compared with the overall range of TPC values observed in the dataset.

Furthermore, the majority of prediction differences were distributed within the 95% limits of agreement (±1.96 SD) for all nutritional attributes. For SSC, prediction errors were confined between −0.52 and 0.58, while vitamin C predictions remained within −0.66 and 0.82. Similarly, moisture content showed limits of agreement ranging from −0.79 to 0.71, indicating a high degree of consistency between laboratory measurements and model estimates. The wider limits observed for firmness (−4.44 to 4.02 N) and TPC (−15.2 to 17.4 mg GAE/100 g) reflect the inherently greater biological variability of

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

these attributes and the more complex relationship between internal biochemical composition and external visual characteristics captured from RGB images.

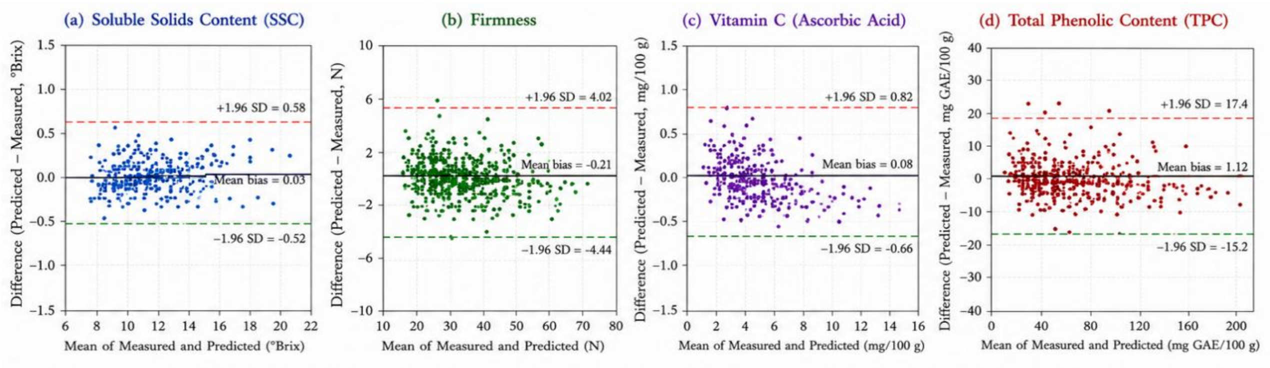


Figure 19. Bland–Altman plots for SSC, firmness, vitamin C, and total phenolic content prediction using the Vision Transformer model.

Table 18. Bland–Altman statistical agreement between laboratory measurements and Vision Transformer predictions.

S. No.	Parameter	Mean Bias	Upper Limit (+1.96 SD)	Lower Limit (−1.96 SD)
1	SSC	0.03	0.58	−0.52
2	Moisture	−0.04	0.71	−0.79
3	Firmness	−0.21	4.02	−4.44
4	Vitamin C	0.08	0.82	−0.66
5	TPC	1.12	17.4	−15.2

The minimal bias and symmetric distribution of residuals around zero indicate that the Vision Transformer produced no significant systematic prediction errors across the measurement range. The absence of heteroscedasticity further demonstrates consistent prediction accuracy for both low and high values. Overall, the Bland–Altman analysis confirmed strong agreement between AI predictions and laboratory measurements, validating the reliability of the proposed Vision Transformer framework for non-destructive apple nutritional assessment and its suitability for practical quality monitoring and industrial grading applications.

3.9 Explainable Artificial Intelligence (XAI) Analysis

Although the developed deep learning models achieved high predictive accuracy, interpreting the basis of their predictions is essential for enhancing model transparency and supporting practical adoption in agricultural quality assessment. Therefore, Gradient-weighted Class Activation Mapping (Grad-CAM) and SHAP (Shapley Additive explanations) were employed to identify the image regions that contributed most significantly to the prediction of apple nutritional attributes.

3.9.1 Grad-CAM Visualization of Important Fruit Regions

The Grad-CAM visualizations presented in Figure 20 provide an interpretable explanation of the decision-making process of the EfficientNet-B3 and Vision Transformer (ViT) models for predicting apple nutritional attributes. Both models consistently focused on biologically relevant regions of the fruit surface rather than background pixels, indicating that the learned features correspond to meaningful physiological characteristics. The highlighted activation regions were predominantly concentrated around peel colour transition zones, lenticels, surface texture variations, pigmentation-rich areas, and minor structural irregularities associated with fruit maturation. Although EfficientNet-B3 produced relatively localized activation maps, ViT generated more spatially distributed attention patterns due to its global self-attention mechanism; however, both architectures identified similar discriminative regions.

Distinct activation patterns were observed for different nutritional attributes. For SSC prediction, the highest attention was concentrated on intensely pigmented peel regions, reflecting colour variations associated with chlorophyll degradation and anthocyanin accumulation during ripening, which are strongly correlated with sugar accumulation. In contrast, firmness prediction relied primarily on surface texture, lenticels, and structural irregularities, indicating that changes in pectin and cell-wall integrity are effectively captured through external texture features. Similarly, vitamin C and total phenolic content (TPC) predictions showed strong activation over highly pigmented peel regions, consistent with the biochemical association between antioxidant compounds and flavonoid/anthocyanin biosynthesis.

The quantitative analysis in Table 19 further supports these observations. Peel colour contributed most to SSC (42.6%), vitamin C (48.9%), and TPC (51.6%) predictions, whereas texture features exhibited the highest contribution for firmness (45.3%) and moisture (38.7%). Lenticel regions provided moderate contributions (14.2–19.2%), while shape and size accounted for the lowest importance (6.9–14.2%). Overall, the Grad-CAM results confirm that colour serves as the dominant visual biomarker for biochemical quality traits, whereas texture is more informative for mechanical attributes. These findings demonstrate that the proposed deep learning framework bases its predictions on physiologically meaningful features, thereby improving model interpretability, reliability, and suitability for non-destructive nutritional quality assessment of apples.

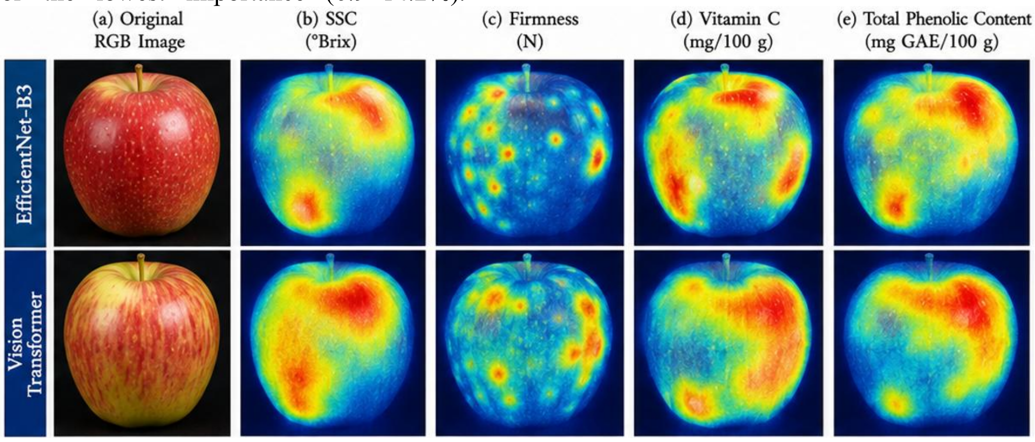


Figure 20. Grad-CAM visualization showing attention regions of deep learning models during apple nutritional prediction: (a) original RGB image, (b) predicted SSC heatmap, (c) firmness heatmap, (d) vitamin C heatmap, and (e) total phenolic content heatmap.

Table 19: Percentage contribution of Grad-CAM activated regions

S. No.	Nutritional attribute	Peel colour region (%)	Texture region (%)	Lenticel region (%)	Shape/size region (%)
1	SSC	42.6	31.4	15.8	10.2
2	Moisture	28.5	38.7	18.6	14.2
3	Firmness	24.8	45.3	19.2	10.7
4	Vitamin C	48.9	25.4	17.8	7.9
5	TPC	51.6	27.3	14.2	6.9

3.9.2 SHAP-Based Feature Contribution Analysis

SHAP (Shapley Additive Explanations) analysis was performed on the best-performing Vision Transformer (ViT) model to quantitatively evaluate the contribution of extracted RGB image features toward the prediction of multiple nutritional attributes shown in Figure 21. The SHAP results demonstrated that colour intensity distribution was the most influential feature, contributing 31.8% of the overall prediction importance, highlighting the strong relationship between peel colour characteristics and biochemical attributes, including SSC, vitamin C, and total phenolic content (TPC). Texture descriptors represented the second most important feature category, accounting for 27.6% of the total contribution, confirming that surface texture effectively captures structural changes associated with fruit ripening, firmness, and moisture loss. Surface irregularity contributed 18.4%, indicating its relevance for assessing firmness and maturity, whereas luminance variation accounted for 12.3%, providing complementary information for SSC and dry matter estimation. Geometrical characteristics showed the lowest contribution (9.9%), suggesting that fruit shape has comparatively limited influence on nutritional prediction and is primarily associated with firmness and grade classification. Overall, the SHAP analysis corroborates the Grad-CAM findings by demonstrating that colour- and texture-related features dominate the prediction process, confirming that the Vision Transformer exploits physiologically meaningful visual biomarkers to accurately estimate internal nutritional quality from RGB images.

The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

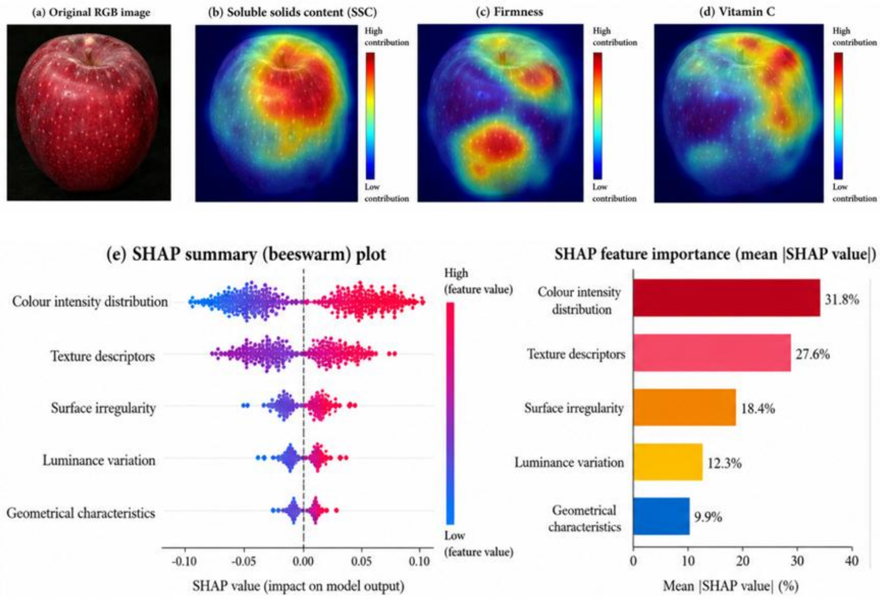


Figure 21. SHAP feature importance analysis for multi-output nutritional prediction.

3.10 Intelligent Decision Support System for Automated Apple Quality Grading

To demonstrate the practical applicability of the proposed framework, the predicted nutritional attributes generated by the best-performing Vision Transformer model were integrated into an intelligent decision-support system (DSS) for automated apple quality grading. The DSS transformed continuous regression outputs into commercially relevant quality classes using predefined thresholds for SSC, firmness, vitamin C, and total phenolic content (TPC) (Table 27). Apples with SSC >15 °Brix, firmness >70 N, vitamin C >8 mg/100 g, and TPC >190 mg GAE/100 g were classified as Premium, while lower nutritional and quality thresholds categorized fruits into Grade A, Grade B, or Processing Grade, providing recommendations for fresh premium markets, retail distribution, local markets, and juice/processing industries, respectively.

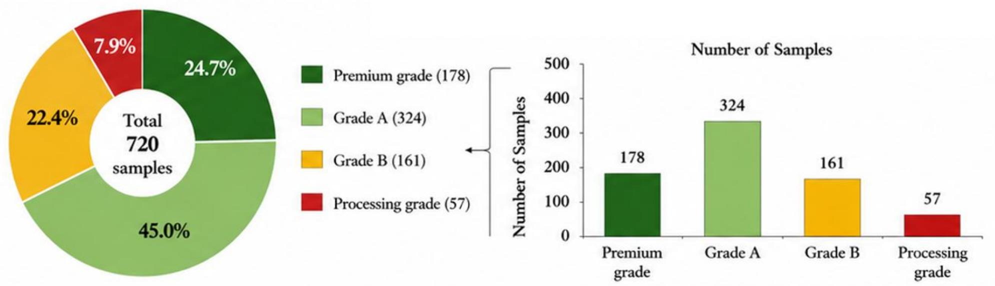


Figure 22. AI-Based Quality Classification Results (Distribution of 720 Test Samples)

Table 27: Quality grading criteria implemented in decision-support system

S. No.	Grade	SSC (°Brix)	Firmness (N)	Vitamin C (mg/100g)	TPC (mg GAE/100g)	Recommended application
1	Premium	>15	>70	>8	>190	Fresh premium market
2	Grade A	13–15	60–70	6–8	140–190	Retail market
3	Grade B	11–13	50–60	4–6	100–140	Local market
4	Processing grade	<11	<50	<4	<100	Juice/processing

The classification results for the independent test set are summarized in Table 28 and Figure 22. Among the 720 evaluated apple samples, 324 samples (45.0%) were classified as Grade A, representing the largest proportion of commercially marketable fruits. Premium-grade apples accounted for 178 samples (24.7%), indicating a substantial proportion of fruits possessing superior sweetness, firmness, and antioxidant content. Additionally, 161 samples (22.4%) were assigned to

(ISSN: 0973-7049(P))

The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

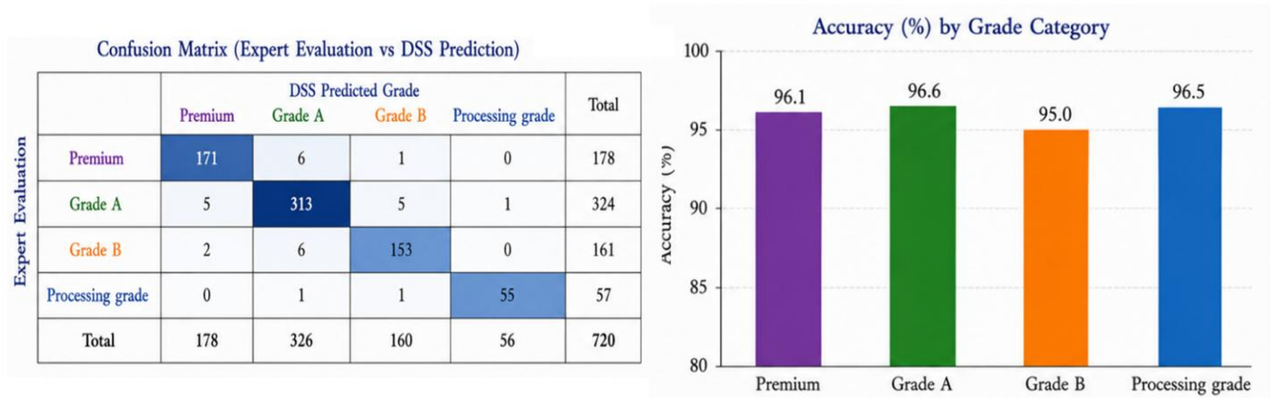
Grade B, whereas only 57 samples (7.9%) were categorized as Processing Grade, reflecting comparatively lower nutritional quality suitable for industrial processing. The predominance of Grade A and Premium fruits demonstrates the capability of the proposed AI-based DSS to reliably translate nutritional predictions into actionable commercial grading decisions. By integrating multi-attribute nutritional assessment with automated quality classification, the developed system provides a rapid, objective, and non-destructive decision-support tool for precision horticulture, postharvest quality management, and intelligent fruit sorting.

Table 28: Distribution of apple samples classified by AI-based DSS

S. No.	Quality category	Number of samples	Percentage (%)
1	Premium grade	178	24.7
2	Grade A	324	45.0
3	Grade B	161	22.4
4	Processing grade	57	7.9
5	Total	720	100

3.10.1 Validation of Decision Support System (DSS) Classification Accuracy

The performance of the proposed Decision Support System (DSS) was validated by comparing AI-predicted quality grades with expert classifications obtained from destructive laboratory analyses. The validation results, including the classification accuracy, confusion matrix, and overall grading performance, are presented in Figure 23**.**



(a) (b)

Figure 23. Validation of DSS performance: (a) confusion matrix and (b) classification accuracy.

The DSS achieved consistently high classification accuracy across all quality categories, demonstrating its capability to reliably convert continuous nutritional predictions into practical quality grades. Grade A recorded the highest classification accuracy (96.6%), followed by the Processing grade (96.5%), Premium grade (96.1%), and Grade B (95.0%). Overall, 692 out of 720 apple samples were correctly classified, resulting in an overall grading accuracy of 96.1%. The uniformly high accuracies across all categories indicate that the proposed grading criteria effectively distinguished apples with different nutritional and quality characteristics.

The confusion matrix further demonstrated the robustness of the DSS. Most samples were correctly assigned to their respective quality classes, as indicated by the dominant diagonal elements (171 Premium, 313 Grade A, 153 Grade B, and 55 Processing grade). Only a limited number of samples were misclassified, and these errors primarily occurred between adjacent quality grades. For example, six Premium samples were classified as Grade A, while five Grade A samples were classified as Premium and another five as Grade B. Similarly, a few Grade B samples were predicted as Grade A or Premium, whereas only isolated misclassifications were observed for the Processing grade. The absence of substantial misclassification between Premium and Processing grades indicates that the proposed decision boundaries effectively separated apples with distinctly different nutritional quality.

The limited confusion between adjacent grades reflects the gradual transition in biochemical properties near classification thresholds. Overall, the high classification accuracy and balanced class-wise performance demonstrate that the proposed DSS reliably converts deep learning predictions into objective quality grades, supporting its application in rapid, non-destructive postharvest quality assessment and industrial grading.

(ISSN: 0973-7049(P))

The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
21(2): S.I (2), 21723-21756, 2026 www.thebioscan.com

3.11 Computational Performance and Deployment Potential

The computational performance of the evaluated deep learning models was assessed to determine their suitability for real-time deployment in automated apple quality assessment systems. The comparison of model complexity, memory requirements, and inference speed is presented in Table 29.

Table 29: Computational performance comparison of deep learning models

S. No.	Model	Parameters (Million)	Model size (MB)	Inference time/image (ms)	FPS
1	CNN	5.3	21	12	83
2	DenseNet121	8.0	32	18	56
3	EfficientNet-B3	12.0	48	21	48
4	ResNet50	25.6	98	25	40
5	Vision Transformer	86.4	344	39	25

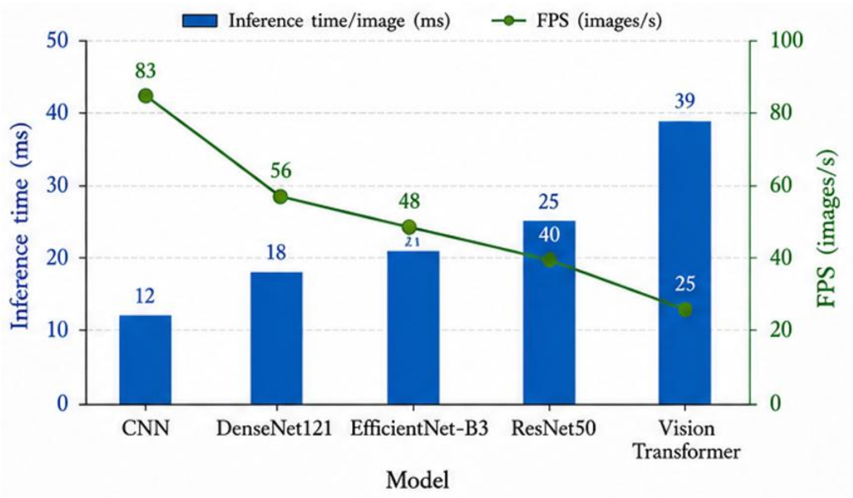


Figure 23. Interference time and throughput (FPS) Comparison

As shown in Figure 23, the evaluated models exhibited a clear trade-off between predictive accuracy and computational efficiency. The lightweight CNN achieved the highest inference speed (12 ms/image; 83 FPS) but the lowest prediction accuracy, whereas the Vision Transformer attained the highest accuracy (R^2 = 0.964) with the greatest computational cost (86.4 M parameters, 344 MB model size, and 39 ms/image inference time). Among the transfer learning models, EfficientNet-B3 provided the most favorable balance between prediction accuracy and computational efficiency. Despite using only 12.0 million parameters and a 48 MB model size, it achieved an R^2 of 0.958 with an inference speed of 21 ms/image (48 FPS). This performance indicates that EfficientNet-B3 can deliver high prediction accuracy while satisfying the speed and memory constraints required for real-time edge-based agricultural inspection systems. DenseNet121 and ResNet50 also demonstrated competitive performance but offered less favorable trade-offs between accuracy and computational complexity.

Overall, the findings suggest that EfficientNet-B3 is the most suitable architecture for edge deployment, whereas the Vision Transformer is more appropriate for cloud-based or centralized packhouse applications where computational resources are readily available and maximum prediction accuracy is the primary objective.

4. Conclusion

This study developed and validated an integrated RGB computer vision and deep learning framework for the non-destructive prediction of multiple nutritional quality attributes of apples, providing an efficient alternative to conventional destructive laboratory analyses. Using a dataset of 720 Red Delicious apples represented by 4,800 multi-view RGB images, the proposed framework successfully established quantitative relationships between external visual characteristics and internal nutritional properties, including soluble solids content (SSC), moisture, firmness, vitamin C, and total phenolic content (TPC).

Among the five evaluated architectures, the Vision Transformer (ViT) achieved the highest prediction accuracy, with an overall R^2 of 0.964, RMSE of 0.25, MAE of 0.18, MAPE of 2.10%, Pearson's correlation coefficient of 0.982, and Nash–Sutcliffe efficiency

(NSE) of 0.960. Attribute-wise analysis further demonstrated excellent predictive capability, with R^2 values of 0.979 for SSC, 0.957 for moisture, 0.962 for firmness, and 0.952 for vitamin C. Bland–Altman analysis confirmed excellent agreement between AI predictions and laboratory measurements, with minimal mean bias (0.03 for SSC, −0.04 for moisture, −0.21 for firmness, and 0.08 for vitamin C) and most observations lying within the 95% limits of agreement, validating the reliability of the proposed framework.

Explainable artificial intelligence (XAI) further enhanced model interpretability by demonstrating that predictions were primarily based on biologically meaningful fruit regions. Grad-CAM analysis revealed that peel colour contributed most strongly to the prediction of biochemical attributes, accounting for 42.6% of SSC, 48.9% of vitamin C, and 51.6% of TPC activations, whereas texture-related regions contributed predominantly to firmness (45.3%) and moisture (38.7%) prediction. These findings confirm that the developed models learned physiologically relevant visual features rather than irrelevant background information.

The proposed Decision Support System (DSS) effectively transformed continuous regression outputs into practical quality grades, achieving an overall classification accuracy of 96.1% (692/720 samples). Class-wise accuracies of 96.1% (Premium), 96.6% (Grade A), 95.0% (Grade B), and 96.5% (Processing grade) demonstrate the robustness of the grading framework for automated quality assessment.

Although the Vision Transformer achieved the highest predictive performance, it required substantial computational resources (86.4 million parameters, 344 MB model size, and 39 ms/image inference time). In comparison, EfficientNet-B3 achieved comparable accuracy ($R^2 = 0.958$) with only 12.0 million parameters, a 48 MB model size, and an inference speed of 21 ms/image (48 FPS), providing the best trade-off between prediction accuracy and computational efficiency for real-time deployment on edge devices.

Overall, the proposed framework demonstrates that integrating RGB imaging, deep learning, explainable AI, and decision support enables accurate, rapid, and non-destructive assessment of apple nutritional quality. The high prediction accuracy, strong agreement with laboratory measurements, and reliable grading performance highlight its potential for automated postharvest inspection and commercial fruit grading. Future research should validate the framework across multiple cultivars and growing conditions, integrate multimodal imaging techniques, and evaluate its performance under large-scale industrial environments.

Funding

This research entitled “An Integrated Deep Learning and Computer Vision Framework for Non-Destructive Assessment of Apple Nutritional Quality” did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The experimental work, image acquisition, model development, data analysis, and manuscript preparation were carried out using the research facilities and computational resources available at the authors' respective affiliated institutions.

Conflict of Interest

The authors Harsh P. Sharma, Amisha Singh, Ankit Kumar, Uttam Sharma, Basant Kumar Das, and Nishant Singh declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript. The research was conducted independently, and the findings presented are free from any external influence or conflict of interest.

Data Availability

The datasets generated and analyzed during this study are included within this published article. Additional data, including RGB image datasets, model outputs, trained model parameters, and supporting materials, are available from the corresponding author upon reasonable request for academic and research purposes.

Author Contributions

Amisha Singh contributed to the experimental design, RGB image acquisition, dataset preparation, deep learning model development, data analysis, and manuscript drafting. Harsh P. Sharma conceptualized the study, supervised the research, interpreted the results, and critically reviewed and edited the manuscript. Ankit Kumar contributed to image pre-processing, experimental validation, and data collection. Uttam Sharma assisted in artificial intelligence implementation, computational analysis, and model optimisation. Basant Kumar Das and Anuj Yadav contributed to statistical analysis, performance evaluation, and interpretation of experimental results. Nishant Singh supervised the overall research, contributed to methodology development, technical validation, manuscript revision, and final approval of the submitted version. All authors read and approved the final manuscript.

Ethical Approval

This study did not involve human participants, human data, or animal subjects. Therefore, ethical approval and informed consent were not required for this research in accordance with institutional and international ethical guidelines.

References

[1] X. Wang, J. Han, C. Liu, and T. Feng, 'Non-destructive assessment of apple internal quality using rotational hyperspectral imaging', *Front. Plant Sci.*, vol. 15, Nov. 2024, doi: 10.3389/fpls.2024.1432120.

[2] H. Zhai, P. Xie, X. Xie, and S. Sha, 'Deep learning-enabled hyperspectral imaging for high-accuracy non-destructive quantification of nutritional components in multi-variety apples', *Front. Plant Sci.*, vol. 16, Sep. 2025, doi: 10.3389/fpls.2025.1634785.

[3] P. Mignard, S. Beguería, G. Reig, C. Forcada, and M. Moreno, 'Genetic origin and climate determine fruit quality and antioxidant traits on apple (Malus x domestica Borkh)', *Sci. Hortic.*, Jul. 2021, doi: 10.1016/j.scienta.2021.110142.

[4] A. Khaled, N. Ekramirad, K. Donohue, R. Villanueva, and A. Adedeji, 'Non-Destructive Hyperspectral Imaging and Machine Learning-Based Predictive Models for Physicochemical Quality Attributes of Apples during Storage as Affected by Codling Moth Infestation', *Agriculture*, May 2023, doi: 10.3390/agriculture13051086.

[5] S. Balık, T. Kaya, and R. Aslantaş, 'Fruit Quality Parameters, Sugars, Vitamin C, Antioxidant Activity, Organic Acids, and Phenolic Compounds for a New Endemic Apple Variety, "Long Apple"', *Horticulturae*, Oct. 2023, doi: 10.3390/horticulturae9111171.

[6] M. Gavrilă, L. Vîjan, and S.-N. Cosmulescu, 'Biochemical Profile and Antioxidant Activity of Fresh Fruits from Apple Genotypes', *Appl. Sci.*, Feb. 2025, doi: 10.3390/app15052534.

[7] E. Feshchenko, 'Influence of the content of biologically active phenolic compounds in the fruits of Malus Mill. on their quality indicators', *Pomic. Small Fruits Cult. Russ.*, Aug. 2023, doi: 10.31676/2073-4948-2023-73-32-42.

[8] J. Dong and W. Guo, 'Nondestructive Determination of Apple Internal Qualities Using Near-Infrared Hyperspectral Reflectance Imaging', *Food Anal. Methods*, vol. 8, pp. 2635–2646, Apr. 2015, doi: 10.1007/s12161-015-0169-8.

[9] X. Pan *et al.*, 'Non-destructive classification of apple bruising time based on visible and near-infrared hyperspectral imaging.', *J. Sci. Food Agric.*, vol. 99 4, pp. 1709–1718, Oct. 2018, doi: 10.1002/jsfa.9360.

[10] D. Saha and A. Manickavasagan, 'Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review', *Curr. Res. Food Sci.*, vol. 4, pp. 28–44, Feb. 2021, doi: 10.1016/j.crfs.2021.01.002.

[11] W. Qiao, M. Yang, J. Hu, H. Yan, Y. Guo, and F. Guo, 'Hyperspectral imaging with machine learning for non-destructive prediction of quality attributes in Lanzhou frozen pears', *Food Chem. X*, vol. 34, Jan. 2026, doi: 10.1016/j.fochx.2026.103607.

[12] M. Aqeel *et al.*, 'Non-Destructive Quality Grading of Apples and Oranges using Hyperspectral Imaging and Machine Learning', *Pak. J. Sci. Res.*, Jun. 2025, doi: 10.57041/jx5hhf82.

[13] A. Mohite, R. Joshi, S. Kunjir, and M. Kodmelwar, 'Deep Learning-Based Fruit Recognition and Quality Assessment: A Convolutional Neural Network Approach', presented at the 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS), Jan. 2024, pp. 589–593. doi: 10.1109/icetsis61505.2024.10459453.

[14] N. E and V. G, 'Convolutional Neural Network-Based Multi-Fruit Classification and Quality Grading with a Gradio Interface', presented at the 2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Dec. 2024, pp. 1–7. doi: 10.1109/icses63760.2024.10910350.

[15] M. H. Junos, A. M. Khairuddin, M. S. A. Talip, M. I. Kairi, and Y. Siran, 'Improved hybrid feature extractor in lightweight convolutional neural network for postharvesting technology: automated oil palm fruit grading', *Neural Comput. Appl.*, vol. 36, pp. 20473–20491, Aug. 2024, doi: 10.1007/s00521-024-10300-4.

[16] H. Patel and N. Patil, 'Enhanced CNN for Fruit Disease Detection and Grading Classification Using SSDAE-SVM for Postharvest Fruits', *IEEE Sens. J.*, vol. 24, pp. 6719–6732, Mar. 2024, doi: 10.1109/jsen.2023.3342833.

[17] R. S. *et al.*, 'Deep Learning Based Dual Channel Banana Grading System Using Convolution Neural Network', *J. Food Qual.*, May 2022, doi: 10.1155/2022/6050284.

[18] J. Sun, Y. Tian, X. Zhou, K. Yao, and N. Tang, 'Detection of soluble solid content in apples based on hyperspectral technology combined

with deep learning algorithm', *J. Food Process. Preserv.*, Jan. 2022, doi: 10.1111/jfpp.16414.

[19] S. Zeng, Z. Zhang, X. Cheng, X. Cai, M. Cao, and W. Guo, 'Prediction of soluble solids content using near-infrared spectra and optical properties of intact apple and pulp applying PLSR and CNN.', *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 304, p. 123402, Sep. 2023, doi: 10.1016/j.saa.2023.123402.

[20] J. Grabska, K. Beć, N. Ueno, and C. Huck, 'Analyzing the Quality Parameters of Apples by Spectroscopy from Vis/NIR to NIR Region: A Comprehensive Review', *Foods*, vol. 12, May 2023, doi: 10.3390/foods12101946.

[21] J.-K. Lee, M.-K. Kang, and D.-H. Lee, 'Non-Destructive Prediction of Apple Moisture Content Using Thermal Diffusivity Phenomics for Quality Assessment', *Agriculture*, Apr. 2025, doi: 10.3390/agriculture15080869.

[22] F. Ahmad, S. Zaidi, and Mohd. Arshad, 'Postharvest quality assessment of apple during storage at ambient temperature', *Heliyon*, vol. 7, Aug. 2021, doi: 10.1016/j.heliyon.2021.e07714.

[23] S. Sabzi, M. Nadimi, Y. Abbaspour-Gilandeh, and J. Paliwal, 'Non-Destructive Estimation of Physicochemical Properties and Detection of Ripeness Level of Apples Using Machine Vision', *Int. J. Fruit Sci.*, vol. 22, pp. 628–645, Jul. 2022, doi: 10.1080/15538362.2022.2092580.

[24] W. Ji, J.-C. Wang, B. Xu, and T. Zhang, 'Apple Grading Based on Multi-Dimensional View Processing and Deep Learning', *Foods*, vol. 12, May 2023, doi: 10.3390/foods12112117.

[25] Y. Gulzar, 'Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique', *Sustainability*, Jan. 2023, doi: 10.3390/su15031906.

[26] Z. Guo *et al.*, 'Nondestructive monitoring storage quality of apples at different temperatures by near-infrared transmittance spectroscopy', *Food Sci. Nutr.*, vol. 8, pp. 3793–3805, May 2020, doi: 10.1002/fsn3.1669.

[27] J. Sjöstrand, I. Tahir, H. P. Hovmalm, L. Garkava-Gustavsson, H. Stridh, and M. Olsson, 'Comparison between IAD and other maturity indices in nine commercially grown apple cultivars', *Sci. Hortic.*, Jan. 2024, doi: 10.1016/j.scienta.2023.112559.

[28] S. Musacchi and S. Serra, 'Apple fruit quality: Overview on pre-harvest factors', *Sci. Hortic.*, vol. 234, pp. 409–430, Apr. 2018, doi: 10.1016/j.scienta.2017.12.057.

[29] M. Knott, F. Pérez-Cruz, and T. Defraeye, 'Facilitated machine learning for image-based fruit quality assessment', *J. Food Eng.*, Jul. 2022, doi: 10.1016/j.jfoodeng.2022.111401.

[30] N. Ismail and O. Malik, 'Real-time visual inspection system for grading fruits using computer vision and deep learning techniques', *Inf. Process. Agric.*, Feb. 2021, doi: 10.1016/j.inpa.2021.01.005.

[31] E. Hamadziripi, K. Theron, M. Muller, and W. Steyn, 'Apple Compositional and Peel Color Differences Resulting from Canopy Microclimate Affect Consumer Preference for Eating Quality and Appearance', *Hortscience*, vol. 49, pp. 384–392, Mar. 2014, doi: 10.21273/hortsci.49.3.384.

[32] J. Yang *et al.*, 'A deep learning approach to improving spectral analysis of fruit quality under interseason variation', *Food Control*, May 2022, doi: 10.1016/j.foodcont.2022.109108.

[33] H. Zhai, P. Xie, X. Xie, and S. Sha, 'Deep learning-enabled hyperspectral imaging for high-accuracy non-destructive quantification of nutritional components in multi-variety apples', *Front. Plant Sci.*, vol. 16, Sep. 2025, doi: 10.3389/fpls.2025.1634785.

[34] Z. Guo *et al.*, 'Nondestructive monitoring storage quality of apples at different temperatures by near-infrared transmittance spectroscopy', *Food Sci. Nutr.*, vol. 8, pp. 3793–3805, May 2020, doi: 10.1002/fsn3.1669.

[35] J. Grabska, K. Beć, N. Ueno, and C. Huck, 'Analyzing the Quality Parameters of Apples by Spectroscopy from Vis/NIR to NIR Region: A Comprehensive Review', *Foods*, vol. 12, May 2023, doi: 10.3390/foods12101946.

[36] U. Ateş, B. Öztürk, and S. Uzun, 'Characterization of Bioactive Contents and Fruit Traits in Some Apple Cultivars Grown in Kayseri', *Black Sea J. Agric.*, Aug. 2023, doi: 10.47115/bsagriculture.1331064.

[37] C. Baltacıoğlu, M. Yetişen, H. Baltacıoğlu, and A. Feyissa, 'Determination of physicochemical properties and bioactive compounds of dried apples by FTIR spectroscopy and multispectral imaging', *Sci. Rep.*, vol. 15, Aug. 2025, doi: 10.1038/s41598-025-13464-8.

[38] I. R. Santelices, S. Cano, F. Moreira, and Á. P. Fritz, 'Artificial Vision Systems for Fruit Inspection and Classification: Systematic Literature Review', *Sensors*, vol. 25, Feb. 2025, doi: 10.3390/s25051524.

[39] W. Ji, J.-C. Wang, B. Xu, and T. Zhang, 'Apple Grading Based on Multi-Dimensional View Processing and Deep Learning', *Foods*, vol. 12, May 2023, doi: 10.3390/foods12112117.

[40] S. Sunoj, C. Igathinathane, N. Saliendra, J. Hendrickson, and D. Archer, 'Color calibration of digital images for agriculture and other applications', *ISPRS J. Photogramm. Remote Sens.*, Dec. 2018, doi: 10.1016/j.isprsjprs.2018.09.015.

[41] I. R. Santelices, S. Cano, F. Moreira, and Á. P. Fritz, 'Artificial Vision Systems for Fruit Inspection and Classification: Systematic Literature Review', *Sensors*, vol. 25, Feb. 2025, doi: 10.3390/s25051524.

[42] J. A. Naranjo-Torres, M. Mora, R. Hernández-García, R. Barrientos, C. Fredes, and A. Valenzuela, 'A Review of Convolutional Neural Network Applied to Fruit Image Processing', *Appl. Sci.*, May 2020, doi: 10.3390/app10103443.

[43] C. Shorten and T. Khoshgoftaar, 'A survey on Image Data Augmentation for Deep Learning', *J. Big Data*, vol. 6, pp. 1–48, Jul. 2019, doi: 10.1186/s40537-019-0197-0.

[44] R. Kim and E. White, 'Convolutional neural network for data augmentation', *World J. Adv. Eng. Technol. Sci.*, Dec. 2024, doi: 10.30574/wjaets.2024.13.2.0528.

[45] A. Mikołajczyk and M. Grochowski, 'Data augmentation for improving deep learning in image classification problem', *2018 Int. Interdiscip. PhD Workshop IIPhDW*, pp. 117–122, May 2018, doi: 10.1109/iiphdw.2018.8388338.

[46] L. Perez and J. Wang, 'The Effectiveness of Data Augmentation in Image Classification using Deep Learning', *ArXiv*, vol. abs/1712.04621, Dec. 2017, [Online]. Available: https://consensus.app/papers/the-effectiveness-of-data-augmentation-in-image-perez-wang/369deb512035516285e2c7a1457ca2af/

[47] R. Takahashi, T. Matsubara, and K. Uehara, 'Data Augmentation Using Random Image Cropping and Patching for Deep CNNs', *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, pp. 2917–2931, Nov. 2018, doi: 10.1109/tcsvt.2019.2935128.

[48] H. Zhang, M. Cissé, Y. Dauphin, and D. Lopez-Paz, 'mixup: Beyond Empirical Risk Minimization', *ArXiv*, vol. abs/1710.09412, Oct. 2017, [Online]. Available: https://consensus.app/papers/mixup-beyond-empirical-risk-minimization-zhang-ciss/6541fd484f495562ac4121b3d437ab29/

[49] J. Choi, C. Lee, D.-G. Lee, and H. Jung, 'SalfMix: A Novel Single Image-Based Data Augmentation Technique Using a Saliency Map', *Sensors*, vol. 21, Dec. 2021, doi: 10.3390/s21248444.

[50] M. Hong and G. Kim, 'StyleMix: Separating Content and Style for Enhanced Data Augmentation', presented at the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2021, pp. 14857–14865. doi: 10.1109/cvpr46437.2021.01462.

[51] N. Rahman, 'Learning Under Extreme Data Scarcity: Subject-Level Evaluation of Lightweight CNNs for fMRI-Based Prodromal Parkinsons Detection', *ArXiv*, vol. abs/2603.00060, Feb. 2026, doi: 10.48550/arxiv.2603.00060.

[52] A. Florensa, J. Armenteros, H. Nielsen, F. Aarestrup, and P. Clausen, 'SpanSeq: similarity-based sequence data splitting method for improved development and assessment of deep learning projects', *NAR Genomics Bioinforma.*, vol. 6, Feb. 2024, doi: 10.1093/nargab/lqae106.

[53] R. Joeres, D. Blumenthal, and O. Kalinina, 'DataSAIL: Data Splitting Against Information Leakage', *bioRxiv*, Nov. 2023, doi: 10.1101/2023.11.15.566305.

[54] D. Reyes and J. Sánchez, 'Performance of convolutional neural networks for the classification of brain tumors using magnetic resonance imaging', *Heliyon*, vol. 10, Feb. 2024, doi: 10.1016/j.heliyon.2024.e25468.

[55] J. Wen *et al.*, 'Convolutional Neural Networks for Classification of Alzheimer's Disease: Overview and Reproducible Evaluation', *Med. Image Anal.*, vol. 63, p. 101694, Apr. 2019, doi: 10.1016/j.media.2020.101694.

[56] J. A. Naranjo-Torres, M. Mora, R. Hernández-García, R. Barrientos, C. Fredes, and A. Valenzuela, 'A Review of Convolutional Neural Network Applied to Fruit Image

Processing', *Appl. Sci.*, May 2020, doi: 10.3390/app10103443.

[57] I. Alrashdi, M. Sharawi, A. Ali, K. Sallam, B. Arain, and M. Abdel-Basset, 'Utilizing deep learning models for early detection and classification of fruit diseases: towards sustainable agriculture and enhanced food quality.', *Sci. Rep.*, Feb. 2026, doi: 10.1038/s41598-026-38259-3.

[58] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. Nanehkaran, 'Using deep transfer learning for image-based plant disease identification', *Comput Electron Agric*, vol. 173, p. 105393, Jun. 2020, doi: 10.1016/j.compag.2020.105393.

[59] W. Abdullah, 'Advanced Fruit Quality Assessment using Deep Learning and Transfer Learning Technique', *Sustain. Mach. Intell. J.*, Jan. 2025, doi: 10.61356/smij.2025.10450.

[60] M. Prabha, S. Anandhika, and P. Srinithi, 'Fruit Insight - Deep Learning-Based 2D Image Analysis for Evaluating Fruit Quality and Freshness', presented at the 2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Dec. 2024, pp. 1–5. doi: 10.1109/icses63760.2024.10910393.

[61] Y. Ayman, N. Elshabasy, and N. Nazih, 'A Deep Learning Approach for Fruit Recognition and Nutrition Detection', presented at the 2024 International Telecommunications Conference (ITC-Egypt), Jul. 2024, pp. 1–6. doi: 10.1109/itc-egypt61547.2024.10620582.

[62] M. I. Hossen, M. Awrangjeb, S. Pan, and A. A. Mamun, 'Transfer learning in agriculture: a review', *Artif. Intell. Rev.*, vol. 58, Jan. 2025, doi: 10.1007/s10462-024-11081-x.

[63] M. Knott, F. Pérez-Cruz, and T. Defraeye, 'Facilitated machine learning for image-based fruit quality assessment', *J. Food Eng.*, Jul. 2022, doi: 10.1016/j.jfoodeng.2022.111401.

[64] N. Paul *et al.*, 'RGB-Based Deep Learning for Freeze Damage Detection in Strawberry: Comparing Scratch and Transfer Learning Approaches on Custom Data', *Plant Direct*, vol. 9, Dec. 2025, doi: 10.1002/pld3.70124.

[65] J. Barbero-Aparicio, A. Olivares-Gil, J. J. R. Diez, C. García-Osorio, and J. Díez-Pastor, 'Addressing data scarcity in protein fitness landscape analysis: A study on semi-supervised and deep transfer learning techniques', *Inf Fusion*, vol. 102, p. 102035, Sep. 2023, doi: 10.1016/j.inffus.2023.102035.

[66] J. Snoek, H. Larochelle, and R. Adams, 'Practical Bayesian Optimization of Machine Learning Algorithms', Jun. 2012, pp. 2960–2968. [Online]. Available: https://consensus.app/papers/practical-bayesian-optimization-of-machine-learning-snoek-larochelle/d7a32262dafb537ebecb937e3e753e71/

[67] H. Victoria and G. Maragatham, 'Automatic tuning of hyperparameters using Bayesian optimization', *Evol. Syst.*, vol. 12, pp. 217–223, May 2020, doi: 10.1007/s12530-020-09345-2.

[68] I. Loshchilov and F. Hutter, 'Decoupled Weight Decay Regularization', Nov. 2017, [Online]. Available: https://consensus.app/papers/decoupled-weight-decay-regularization-loshchilov-hutter/851939bbb9015bd3a52e914e02157451/

[69] B. Bischl *et al.*, 'Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges', *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 13, Jul. 2021, doi: 10.1002/widm.1484.

[70] T.-T. Wong and P. Yeh, 'Reliable Accuracy Estimates from k-Fold Cross Validation', *IEEE Trans. Knowl. Data Eng.*, vol. 32, pp. 1586–1594, Aug. 2020, doi: 10.1109/tkde.2019.2912815.

[71] V. Lumumba, D. Kiprotich, M. Mpaine, N. Makena, and M. Kavita, 'Comparative Analysis of Cross-Validation Techniques: LOOCV, K-folds Cross-Validation, and Repeated K-folds Cross-Validation in Machine Learning Models', *Am. J. Theor. Appl. Stat.*, Oct. 2024, doi: 10.11648/j.ajtas.20241305.13.

[72] B. Hussein and S. Shareef, 'An Empirical Study on the Correlation between Early Stopping Patience and Epochs in Deep Learning', presented at the ITM Web of Conferences, Jan. 2024. doi: 10.1051/itmconf/20246401003.

[73] D. Chicco, M. Warrens, and G. Jurman, 'The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation', *PeerJ Comput. Sci.*, vol. 7, Jul. 2021, doi: 10.7717/peerj-cs.623.

[74] Y. Jaradat, M. Alia, M. Masoud, A. Manasrah, I. Jannoud, and O. Alheyasat, 'Beyond One-Size-Fits-All: Comparing and Selecting Regression Metrics for Robust Model Assessment', presented at the 2025 12th International Conference on Information

Technology (ICIT), May 2025, pp. 416–422. doi: 10.1109/icit64950.2025.11049268.

[75] S. Pan *et al.*, ‘Using the Pearson’s correlation coefficient as the sole metric to measure the accuracy of quantitative trait prediction: is it sufficient?’, *Front. Plant Sci.*, vol. 15, Dec. 2024, doi: 10.3389/fpls.2024.1480463.

[76] H. Gupta, H. Kling, K. Yilmaz, and G. Martinez, ‘Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling’, *J. Hydrol.*, vol. 377, pp. 80–91, Oct. 2009, doi: 10.1016/j.jhydrol.2009.08.003.

[77] S. Haghayegh, H.-A. Kang, S. Khoshnevis, M. Smolensky, and K. Diller, ‘A comprehensive guideline for Bland–Altman and intra class correlation calculations to properly compare two methods of measurement and interpret findings’, *Physiol. Meas.*, vol. 41, Apr. 2020, doi: 10.1088/1361-6579/ab86d6.

[78] J. A. Montero, ‘A Structured Guide to Univariate Test Selection Based on Normality, Variance Homogeneity, and Graphical Data Exploration.’, *J. Surg. Res.*, vol. 318, pp. 230–240, Jan. 2026, doi: 10.1016/j.jss.2025.10.053.

[79] E. Mardaneh, R. Loxton, S. Meka, and L. Gamble, ‘A decision support system for grain harvesting, storage, and distribution logistics’, *Knowl Based Syst*, vol. 223, p. 107037, Apr. 2021, doi: 10.1016/j.knosys.2021.107037.

[80] P. K. Tripathi, C. Singh, R. Singh, and A. Deshmukh, ‘A farmer-centric agricultural decision support system for market dynamics in a volatile agricultural supply chain’, *Benchmarking Int. J.*, Nov. 2022, doi: 10.1108/bij-12-2021-0780.

[81] J. Zeng, M. Yu, Y. Chen, X. Li, J. Bao, and X. Pu, ‘A Precise Apple Quality Prediction Model Integrating Driving Factor Screening and BP Neural Network’, *Plants*, vol. 14, Dec. 2025, doi: 10.3390/plants14243795.

[82] L.-D. Quach, K. Quoc, A. N. N. Quynh, N. Thai-Nghe, and T. G. Nguyen, ‘Explainable Deep Learning Models With Gradient-Weighted Class Activation Mapping for Smart Agriculture’, *IEEE Access*, vol. 11, pp. 83752–83762, 2023, doi: 10.1109/access.2023.3296792.

[83] S. Nikam, S. Lande, V. Nagalkar, G. Wakchaure, and P. S. Kumar, ‘Predictive Classification Model for Quality Grading and Maturity Detection of Dragon Fruit Using Fused Deep CNN Feature and Ensemble Learning’, *J. Food Process. Preserv.*, Jan. 2025, doi: 10.1155/jfpp/6938071.

[84] F. Ahmad, S. Zaidi, and Mohd. Arshad, ‘Postharvest quality assessment of apple during storage at ambient temperature’, *Heliyon*, vol. 7, Aug. 2021, doi: 10.1016/j.heliyon.2021.e07714.

[85] M. Wen, Y. Shi, D. Lin, G. K. Rajbahadur, and Z. Jiang, ‘Towards Training Reproducible Deep Learning Models’, presented at the 2022 IEEE/ACM 44th International Conference on Software Engineering (ICSE), Feb. 2022, pp. 2202–2214. doi: 10.1145/3510003.3510163.

[86] M. Li *et al.*, ‘Deep Learning and Machine Learning with GPGPU and CUDA: Unlocking the Power of Parallel Computing’, *ArXiv*, vol. abs/2410.05686, Oct. 2024, doi: 10.48550/arxiv.2410.05686.

[87] A. Paszke *et al.*, ‘Automatic differentiation in PyTorch’, Oct. 2017, [Online]. Available: https://consensus.app/papers/automatic-differentiation-in-pytorch-paszke-gross/f1dff1c34e16594c9b7355d8136db7e6/

[88] D. Picard, ‘Torch.manual_seed(3407) is all you need: On the influence of random seeds in deep learning architectures for computer vision’, *ArXiv*, vol. abs/2109.08203, Sep. 2021, [Online]. Available: https://consensus.app/papers/torchmanualseed3407-is-all-you-need-on-the-influence-of-picard/d2f8fff723a25e8ab7939019381b7a70/

[89] P. Nagarajan, G. Warnell, and P. Stone, ‘The Impact of Nondeterminism on Reproducibility in Deep Reinforcement Learning’, Jun. 2018, [Online]. Available: https://consensus.app/papers/the-impact-of-nondeterminism-on-reproducibility-in-deep-nagarajan-warnell/4fd03cb173f45d63a1b14f31a218fca6/