A Computational Framework for Apple Detection Using Fuzzy Logic and Structural Cues

Kai Siong Yow[*]

Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Malaysia

[*] Correspondence: Kai Siong Yow (ksyow@upm.edu.my)

Abstract: Accurate and reliable detection of apples in complex orchard environments remains a challenging task due to varying illumination, cluttering backgrounds, and overlapping fruits. In this paper, the difficulties were tackled with a novel edge-enhanced detection framework proposed to integrate dynamic image smoothing, entropy-based edge amplification, and directional energy-driven contour extraction. An adaptive smoothing filter was adopted with a sigmoid-based weighting function to selectively preserve edge structures while suppressing noise in homogeneous regions. The input of Red Green Blue (RGB) image was subsequently transformed into the Hue, Saturation, and Value (HSV) color space to exploit hue information, thereby improving color-based feature discrimination. The introduction of a hybrid entropy-weighted gradient scheme helped strengthen edge detection, that is, the local image entropy modulated gradient magnitudes to emphasize structured regions. A global threshold was then applied to refine the enhanced edge map. Ultimately, continuous apple contours were extracted using a direction-constrained energy propagation approach, in which connected edge pixels were traced according to compass orientations, thus ensuring accurate contour assembly even under occlusion or low contrast. Experimental evaluations confirmed that the proposed framework substantially improved the accuracy of boundary detection across diverse imaging conditions; its potential application in automated fruit detection and precision harvesting was therefore highlighted.

Keywords: Apple detection; Adaptive smoothing; Entropy-weighted gradient; Edge enhancement; Directional energy contour tracing

1 Introduction

Image processing is a rapidly evolving field that plays a vital role in numerous applications, including medical imaging, remote sensing, surveillance, and computer vision. It encompasses a diverse set of methodologies developed to improve and examine, segment, and interpret visual information from digital images. Traditional image processing methods often struggle with challenges, such as intensity inhomogeneity, noise, and low contrast, which limit their effectiveness in real-world scenarios. To address these challenges, scholars have proposed high technologies in models and algorithms that deploy region-based segmentation, active contour models, and deep learning. These methods enhance the meticulousness and durability of the image processing, especially when there is a complicated environment. The introduction of Graphics Processing Units (GPU) optimized computation has greatly increased the speed of processing related to deep learning-based methods, thus enabling the application of real time more easily. The combination of statistical modeling, convolutional processing, and artificial intelligence persists to spur innovations and provides more accurate and less energy-consuming solutions to various imaging issues [1–5].

Based on the progress of the image processing, recent research activities have engaged in the use of incredible image smoothing algorithms to increase the quality of the image formation and at the same time it retains the vital edge information by giving a fair amount of discussion. Common problems with traditional smoothing techniques include smoothing of important structures or the presence of stair case artifacts. In order to deal with these issues, new techniques are introduced such as edge-aware smoothing model and gradient reconstructions. To take an example, Zeng et al. [6] suggested the weighted sparse gradient reconstruction, which keeps its sharp edges and smooths the flat areas efficiently. This was further elaborated by Matsuoka and Okuda [7], who reduced the gradient0 in the objective smaller stair casing effects and included new constraints to address this issue and resulted in increased robustness. Yada and Sarawadekar [8] were able to formulate multiple-scale edge-smoothing filter in the context

https://doi.org/10.56578/ida040205

of the dehazing of images, which can adapt to different densities of the haze, hence showing better visual clarity. Besides, Al-Ameen [9] proposed an algorithm based on directional variance specifically dedicated to digital image smoothing which softly suppresses noise with a directional awareness.

Expanding on the foundations of image smoothing and segmentation [10, 11], the field of object detection in remote sensing and agricultural imagery has seen substantial growth through the integration of deep learning techniques. Researchers have increasingly turned to convolutional neural networks and object detection frameworks like You Only Look Once (YOLO) to address challenges such as small object detection, varying lighting conditions, and complex backgrounds. Bai et al. [12] and Yi et al. [13] proposed significant improvements in remote sensing applications by reviewing and enhancing YOLO-based models for small object detection. Similarly, Zhang et al. [14] and Wang et al. [15] introduced customized YOLOv5-based frameworks for greenhouse tomato and apple fruit detection, respectively; their studies demonstrated the effectiveness of the frameworks in structured agricultural environments. In addition, Kılıçarslan et al. [16] employed hybrid transfer learning and multi-level feature extraction techniques to distinguish apple varieties with high accuracy. Despite the remarkable progress achieved through the above deep learning-based object detection and classification methods, several limitations persist. Many YOLO-based models, while efficient, often struggle with accurately detecting very small or densely clustered objects in high-resolution remote sensing images, thus leading to missed detection or false positives. In agricultural scenarios, these models could be sensitive to occlusions, overlapping fruits, and seasonal variations in appearance, which may degrade their abilities of generalization across different datasets. Furthermore, the reliance on large annotated datasets poses a challenge in domains where labeled data are scarce or labor-intensive to obtain. From a computational standpoint, the deployment of deep models in real-time or resource-constrained environments remains difficult, especially when model complexity increases. These challenges underline the need for robust and adaptive segmentation as well as smoothing techniques.

In this research paper, we proposed an entropy-weighted gradient and directional edge-enhancement-based model for accurate apple detection in natural orchard scenes. The model functioned through a structured pipeline beginning with adaptive smoothing to suppress background noise while retaining essential image features. It then calculated entropy-weighted gradients to emphasize regions with high information content, particularly around the edges of the object [17, 18]. These enhanced gradients were processed through a directional edge amplification mechanism to strengthen the continuity of the apple boundaries. Finally, an energy-based contour tracing strategy was applied, leveraging both local edge strength and orientation to delineate apple regions accurately, in spite of the occurrence of occlusions, overlapping branches, and varying lighting conditions.

Unlike traditional methods or popular deep learning frameworks like YOLO, which often require extensive labeled datasets and computational resources, our model emphasizes interpretability and data efficiency. By integrating entropy-guided edge analysis with directional reinforcement and fuzzy-driven energy minimization, we provide a rule-based mechanism that performs competitively in challenging real-world orchard environments without requiring massive training efforts. This makes the proposed model especially suitable for low-resource settings where annotated data and GPU support may be limited. Furthermore, YOLO-based models may struggle with occlusions and non-uniform illumination unless retrained on orchard-specific datasets, whereas our approach explicitly incorporates these structural and contextual variances through handcrafted entropy and edge metrics.

The novel contributions of the proposed model are summarized as follows:

- **Entropy-weighted gradient computation:** A new handcrafted approach that assigns local entropy as a weight to gradient magnitudes, is introduced to enable context-aware edge emphasis. Unlike previous works that apply uniform edge enhancement or use fixed weighting schemes, our entropy-guided computation dynamically adapts to local texture variations, hence providing superior differentiation of object boundaries in cluttered backgrounds.

- **Directional edge enhancement mechanism:** A novel directional filtering technique is proposed to enhance edge continuity along dominant orientations. While direction-aware edge processing is not new, our model uniquely combines it with entropy-weighted cues to retain boundary consistency under heavy occlusion and illumination noise; this creates a scenario in which traditional methods and even some Convolutional Neural Network (CNN) may underperform.

- **Energy-based contour tracing algorithm:** A robust contour detection strategy is formulated with a fuzzy energy minimization principle, which balances local edge strength and gradient directionality. This differs from classical snakes or graph-cut methods by integrating contextual entropy cues and fuzzy logic rules, resulting in high boundary adherence with minimal false positives.

- **Fuzzy feature extraction:** A fuzzy logic-based mechanism is employed to extract uncertainty-aware features that capture both intensity variation and contextual texture. This enhances the ability of the model to discern subtle boundaries and structure in low-contrast regions. By leveraging fuzzy membership functions, the model ensures robustness to noise and illumination changes, thus outperforming crisp thresholding or hard segmentation techniques.

- **Computational efficiency and scalability:** The model is lightweight, interpretable, and highly parallelizable. Compared to deep models like YOLOv5 that require pre-trained weights and fine-tuning on apple datasets, our approach runs in real time with minimal memory usage and achieves comparable segmentation accuracy without deep learning infrastructure.

2 Literature Review

In recent years, several studies have explored entropy-based and learning-driven image seg-mentation and enhancement methods, thus contributing significantly to applications in biomedical imaging, agriculture, and object grading.

Gupta et al. [19] proposed a deep learning-enhanced automated mitochondrial segmentation framework for Focused Ion Beam Scanning Electron Microscope (FIB-SEM) images using an entropy-weighted ensemble of multiple convolutional neural networks. The approach assigns adaptive weights based on entropy to improve segmentation confidence and accuracy. This method achieved high precision in segmenting complex cellular structures. Nonetheless, the primary drawback of the model is its high computational complexity and the need to consider extensive and high-quality annotated data sets, which would be inapplicable in real-time processing or environment where computer resources are insufficient.

Gill and Khehra [20] also proposed a segmentation of apple image using Teaching Learning Based Optimization (TLBO) with minimum cross entropy thresholding. The model is able to partition boundaries of an apple satisfactorily even in different lighting situations. It is unique because it is optimized to choose the threshold and makes it robust to variations in image brightness. Although this model works well, it could falsely recognize objects or noise around the apple as an apple, in particular when the apple overlaps or there are some occluding elements. The process of optimization is computationally costly, thus impairing its real-time applicability.

According to Wang et al. [21], multi-featured grading model of apples combined the benefits of entropy-based weighting-mechanism and multi-layered perceptron (MLP). The model densely extracted and incorporated several features including color, texture, and shape; apple grades were then classified with an MLP. Although the hybrid structure enhanced grading error rates in terms of different lighting conditions, surface features, leaves, stems, and close fruits, some misclassifications were presented as part of the non-apple features which are similar to those of the apple. Moreover, the model depended on handcrafted feature extraction, which prevents its generalization to unseen apple varieties and backgrounds.

All these studies demonstrated the opportunities of combining entropy and learning so as to build a metal framework for better image processing. Despite satisfactory results in performance, the drawbacks of the combination include complexity of calculations, misidentification of non-targeted items, and over-reliance on manually crafted features; all these require more robust and adaptable solutions in real-world scenarios.

3 Proposed Model

On the basis of the advantages and identified deficiencies of the models previously examined, a new entropy-based framework of image segmentation was proposed and implemented in this study to support robust and accurate detection of objects, especially apple objects, by bounding boxes. The proffered model incorporated an entropy-weighted feature separation unit of the kind with the help of a purified module. It also provided fuzzy energy that was functional to improve the precision of the segmentation with a reduction in the amount of confusing non-targets, which could be leaves, branches, and clusters of fruits. As illustrated in Figure 1, by leveraging adaptive local entropy and statistical similarity, the model could accurately delineate object boundaries and generate reliable bounding boxes even under noisy, low-contrast, and occluded conditions. Furthermore, a lightweight computational structure was adopted to ensure rapid inference and enabled the model to function effectively in real-time farming scenarios and smart farming systems.

3.1 Preprocessing and Smoothing

Let the input RGB image be denoted by: $I(x,y) \in \mathbb{R}^{M \times N \times 3}$ where and represent the spatial dimensions, height and width, of the image, and the third dimension corresponds to the Red, Green, and Blue (RGB) color channels. This image undergoes a two-step preprocessing procedure: (i) adaptive smoothing; and (ii) color space transformation.

3.1.1 Adaptive smoothing function

Instead of applying a traditional Gaussian filter, which tends to blur edges and fine details, we introduce a nonlinear adaptive smoothing mechanism that selectively smooths homogeneous regions while preserving edge information. The smoothed image is computed as:

$$I_s(x,y) = \frac{1}{K(x,y)} \sum_{(m,n) \in \Omega} I(x+m, y+n) \cdot \phi(\Delta I(m,n)) \tag{1}$$

where, is a local neighborhood window centered around pixel, typically a square region (e.g., or).

The term represents the absolute intensity difference between the central pixel and its neighboring pixel within the window:

$$\Delta I(m,n) = |I(x,y) - I(x+m, y+n)|$$

To control the contribution of each neighboring pixel based on its similarity to the center pixel, we introduce a sigmoid weighting function:

$$\phi(\Delta I) = \frac{1}{1 + e^{\alpha(\Delta I - T)}}$$

Here, control the sharpness of the transition in the sigmoid function, higher values lead to more selective filtering, a threshold that defines the sensitivity to intensity differences. Pixels with intensity values close to the center will yield higher weights while those with large differences, likely representing edges, will have lower weights.

To normalize the result and ensure proper scaling, we use a normalization factor:

$$K(x,y) = \sum_{(i,j) \in \Omega} \phi(\Delta I(i,j)) \tag{2}$$

This formulation allows the filter to perform edge-preserving smoothing: flat (low-texture) regions are smoothed aggressively while edges and boundaries are preserved, due to their larger intensity variations.
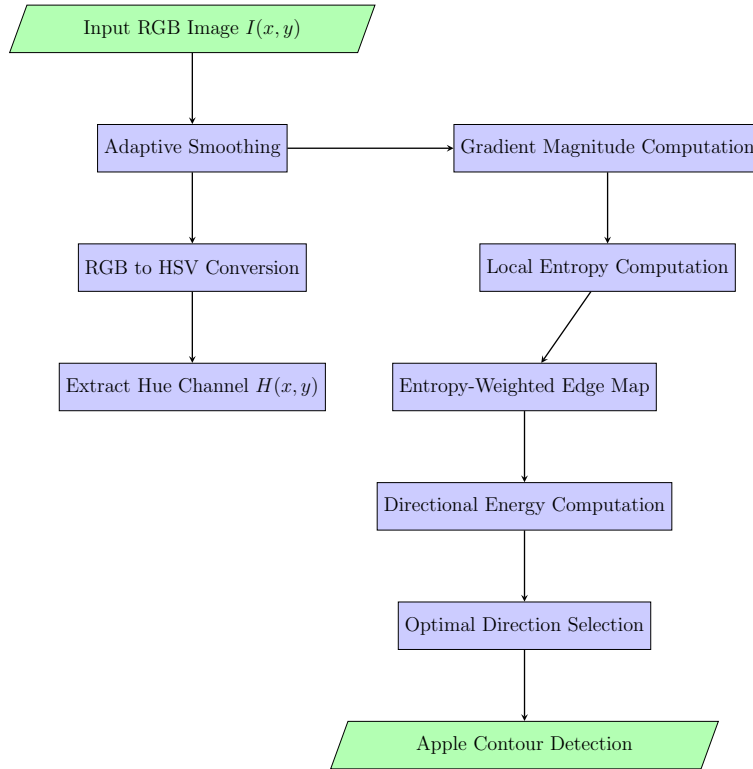

Figure 1. Workflow diagram of the proposed entropy-weighted directional energy model for apple contour detection

3.1.2 Color space transformation

After applying the adaptive smoothing filter, the resulting image was converted from the RGB color space to the Hue-Saturation-Value (HSV) color space, which separated chromatic content (hue) from intensity and colorfulness. This transformation facilitated tasks like segmentation, edge detection, and classification, as hue is often more robust to lighting variations. The transformation is expressed as:

$$HSV(x,y) = \mathrm{rgb\,2\,hsv}\,(I_s(x,y)) \tag{3}$$

Based on the HSV representation, we can derive the hue channel, which captures the prevailing color wavelength of wavelength per pixel:

$$H(x, y) = HSV_H(x, y)$$

The information of this hue is applied later in feature extraction or classification. It will fall under the specific application.

The edges need to be detected in the image with improved tolerance of noise and lighting variations that should be detected by the above algorithm. We used a fusion method which provides both gradient information and local entropy. This fusion would exaggerate non-weak edges with respect to both gradient and structure, being complex in the distribution of local intensity.

3.2 Gradient Magnitude and Local Entropy Computation

As the first step in extracting relevant features of an image, we first computed the gradient magnitude with different intensities at every pixel point and reflected the intensity variation strength. The areas that had sharp sides were emphasized. The gradient magnitude is mathematically expressed as:

$$G(x, y) = \sqrt{I_x(x, y)^2 + I_y(x, y)^2} \tag{4}$$

where, $I_x(x, y)$ and $I_y(x, y)$ denote the partial derivatives of the image intensity I in the horizontal and vertical directions, respectively. These derivatives are commonly approximated using Sobel filters, which convolve the input image with predefined kernel masks designed to emphasize horizontal and vertical intensity transitions.

Once the gradient information was obtained, we proceeded to estimate the local entropy within a neighborhood around each pixel. Entropy is a statistical measure that quantifies the degree of randomness or the complexity of intensity values, which are often used to detect textured or information-rich regions in an image. The value of local entropy at the position $(x_m, y - n)$ is calculated using the following expression:

$$H(x_m, y_n) = -\sum_{k=0}^{L-1} p_k(x_m, y_n) \log(p_k(x_m, y_n) + \epsilon) \tag{5}$$

where:
- $p_k(x, y)$ is the normalized histogram (i.e., probability distribution) of intensity level k within a local window centered at pixel (x, y),
- L is the total number of possible intensity levels in the image (e.g., $L = 256$ for 8-bit grayscale images),
- ϵ is a small positive constant (e.g., $\epsilon = 10^{-8}$) added to prevent the logarithm from becoming undefined when $p_k(x, y) = 0$.

A higher entropy value $H(x, y)$ typically corresponds to regions with greater intensity variation, hence indicating the presence of textures, edges, and object boundaries. This metric is especially useful in tasks like image segmentation, where the distinction between homogeneous and heterogeneous regions is crucial.

3.3 Entropy-Weighted Edge Enhancement for Apple Detection

Accurate edge extraction is essential for distinguishing apple boundaries from complex orchard backgrounds. We combine gradient magnitude with local entropy to form an entropy-weighted edge strength measure to improve robustness. In this context, entropy, complementing the local intensity variation captured by the gradient magnitude, is used as a measure of uncertainty or textural richness. The entropy-weighted formulation is expressed as:

$$E_{\text{custom}}(x, y) = G(x, y) \cdot \left(\frac{H(x, y)}{\max H}\right)^{\beta} \tag{6}$$

where:
- $G(x, y)$ denotes the gradient magnitude at pixel (x, y), typically computed using Sobel or Prewitt operators as $G(x, y) = \sqrt{(I_x)^2 + (I_y)^2}$, where I_x, I_y are partial derivatives of the image intensity,
- $H(x, y)$ represents the local entropy, computed over an $m \times m$ neighborhood window centered at (x, y). Formally, if p_i denotes the probability of gray-level i in the neighborhood, then

$$H(x, y) = -\sum_{i=1}^{L} p_i \log(p_i),$$

where, L is the number of gray levels. The window size m balances locality and stability: smaller windows (e.g., $m = 5$) capture fine details, while larger windows (e.g., $m = 15$) provide robustness against noise. In this work, $m = 9$ was empirically chosen as a trade-off between precision and stability.

- max H is the maximum entropy value observed across the image, serving as a normalization factor to ensure that $H(x,y)/\max H \in [0,1]$,
- β is a tunable exponent (e.g., $\beta = 1.5$) that regulates the contribution of entropy; higher β emphasizes texture-rich areas, while $\beta = 1$ yields a linear influence.

This theoretically grounded formulation allows regions of high uncertainty; for instance, apple surfaces with reflections, blemishes, and fine textures to be emphasized, while suppressing responses in smoother regions such as sky or soil. Therefore, apple contours remain stable against illumination variation and background clutter.

To obtain a binary edge representation, a global threshold τ_E is applied. For reproducibility, τ_E can be determined adaptively, e.g., by Otsu's method, which minimizes intra-class variance, rather than being set manually:

$$E_B(x,y) = \begin{cases} 1, & \text{if } E_{\text{custom}}(x,y) \geq \tau_E \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

This step ensures a principled extraction of edges and yields a clean structural outline of apples to serve as a foundation for subsequent contour-based segmentation, object detection, and quality assessment.

3.4 Directional Energy Computation for Contour Tracing

After generating the binary edge map $E_B(x,y)$, the next step involves extracting the contours of objects such as apples by tracing the connected edge pixels. This is achieved by analyzing the directional energy propagation, which refers to the presence of edge pixels in specific directions around a given location.

For each edge pixel $(x,y) \in E_B$, we examined its 8-connected neighbors in standard compass directions: North (N), North-East (NE), East (E), South-East (SE), South (S), South-West (SW), West (W), and North-West (NW). The directional energy in a given direction θ is defined as:

$$\mathcal{D}_\theta(x,y) = E_B\left(x + \delta_x^\theta, y + \delta_y^\theta\right) \tag{8}$$

where, δ_x^θ and δ_y^θ are the direction-specific offsets associated with direction θ. These offsets guide the traversal to neighboring pixels; for instance:
- $\theta = \text{E} \Rightarrow (\delta_x, \delta_y) = (1,0)$
- $\theta = \text{NE} \Rightarrow (\delta_x, \delta_y) = (1,-1)$
- $\theta = \text{S} \Rightarrow (\delta_x, \delta_y) = (0,1)$, etc.

Mathematically, the continuity of the contour can be ensured by maximizing a connectivity functional:

$$\theta^*(x,y) = \arg\max_\theta \left[\mathcal{D}_\theta(x,y) \cdot w_\theta\right],$$

where, $\mathcal{D}_\theta(x,y) \in \{0,1\}$ and w_θ is a directional weight. For isotropic tracing, $w_\theta = 1$. For directional smoothing, weights can be biased according to the gradient orientation at (x,y), ensuring that tracing aligns with the dominant edge direction. This formalization makes the method reproducible and adaptable to other datasets.

The pixel in the optimal direction $\theta^*(x,y)$ is then selected as the next point on the contour. This process is iteratively applied until a closed loop is formed as in the case of closed objects like apples or until a stopping condition is met.

3.5 Bounding Box Generation for Object Localization

Once the complete contour C_k of an object, e.g., apple, is extracted, we compute the axis-aligned bounding box (AABB) that minimally encloses the contour:

$$B_k = \left(\min_x C_k, \min_y C_k, \max_x C_k, \max_y C_k\right).$$

This definition ensures rigor by providing a mathematically minimal enclosing rectangle. Bounding boxes are widely used in object detection benchmarks, thus allowing direct comparison with standard metrics such as Intersection over Union (IoU). In the context of apple detection, bounding boxes localize individual apples so as to aid in automated harvesting, grading, and yield estimating.

3.6 Fuzzy Feature Extraction

Feature extraction using fuzzy logic is employed to support intelligent classification of apples within detected bounding boxes. Unlike crisp thresholds, fuzzy sets accommodate uncertainty and natural variability in color and shape. The process begins by extracting key features of each candidate region, including hue (color information), geometric descriptors like aspect, ratio, and roundness as well as edge compactness. These features are normalized into the interval [0,1] before fuzzification.

Formally, a fuzzy set F over a feature space X is defined as:

$$F = \{(x, \mu_F(x)) \mid x \in X, \mu_F(x) \in [0,1]\},$$

where, $\mu_F(x)$ is the membership function that maps a feature value x to a degree of membership. For hue-based detection of red apples, we define:

$$\mu_{\text{red}}(H) = \begin{cases} 1, & H < 0.05 \\ \frac{0.1 - H}{0.05}, & 0.05 \leq H < 0.1 \\ 0, & H \geq 0.1 \end{cases} \tag{9}$$

which is a piecewise linear triangular membership function. Unlike the earlier simplified version, this formulation ensures continuity and smooth transitions. Membership functions for shape descriptors such as roundness and elongation, are similarly defined using normalized geometric measures. Visualizations of these functions are provided in Figure 2 to confirm validity and interpretability.

Figure 2. Sample images of apples used to assess the effectiveness of the suggested fuzzy logic-based detection model

The fuzzy inference rules are expressed in the standard Mamdani form. While initial rules are handcrafted based on domain expertise, reproducibility is ensured by validating them against labeled datasets. Moreover, rule optimization can be achieved through adaptive techniques such as genetic algorithms or neuro-fuzzy learning, which tune membership parameters and rule weights based on performance metrics such as accuracy or F1-score. This establishes a pathway for generalization beyond handcrafted rules.

For example:
- **Rule 1**: IF hue is *red* AND shape is round, THEN object is classified as *apple*.
- **Rule 2**: IF hue is *not red* OR shape is elongated, THEN object is classified as *not apple*.

These rules allow flexible decision-making. In practice, the fuzzy inference system aggregates rule outputs using max-min composition and defuzzifies the result into a binary apple/non-apple decision. This mathematically grounded pipeline reduces false positives; for example, red leaves and tomatoes while ensuring reproducibility and robustness across datasets.

4 Discussion and Results

This section shows the results of the experiments and the detailed description of the proposed fuzzy logic-based apple detection model. The model was tested on different types of images of real-world apples (refers to Figure 2) to determine the performance of the model in locating apples with bounding box and feature extraction. The

assessment was based on particular features, such as the accuracy of detection working under different lighting and background conditions, and the model capability to identify an apple object among the similarities in terms of color or shape properties. The experiments show that the combination of characteristics such as color in the hue domain with shape descriptors within the fuzzy inference system has a great impact on the improvement of the accuracy of the detection, especially in tricky orchard images.

The experiment set contained 300 real orchard images, which were retrieved publicly by the MinneApple dataset, in the field scenes. MinneApple dataset served with high spatial resolution apple orchard images and extensive annotations was captured in different occlusions and lighting conditions, so it was a good choice to test apple detectors. The images were all manually checked and extra field images were included to add a variety. The process of annotations was done in the PASCAL VOC scheme and bounding boxes were used to identify instances of the apples.

All experiments and image processing tasks were conducted using MATLAB R2019b, a high-level technical computing environment widely used for image analysis, visualization, and algorithm development. MATLAB offers built-in functions and toolboxes that facilitate efficient matrix manipulation, filtering operations, and edge detection techniques. The software was run on a Windows 10 (64-bit) operating system, equipped with an Intel Core i7 processor and 8 Gigabyte (GB) random access memory (RAM), to ensure the smooth execution of all algorithms, including those requiring substantial computational resources such as entropy-based filtering and gradient computations.

All input images were resized to a fixed resolution of 255×255 pixels using the imresize function to guarantee uniformity and reduce computational complexity. This standardization allows consistent spatial analysis and facilitates a fair comparison across different image samples, particularly during feature extraction and contour detection stages.

We employed a 5-fold cross-validation approach to assess the generalizability of the model. The dataset was randomly split into 80% of training and 20% of testing subsets in each fold. The average precision, recall, and F1-score were computed across all folds to ensure statistical robustness and minimize overfitting. This experimental design supports reproducibility and provides a comprehensive overview of the effectiveness of the model.

To optimize the performance of the proposed apple detection framework, we empirically selected the parameter values listed in Table 1 These values were determined through extensive cross-validation and visual analysis across diverse image samples. The selected configuration strikes a balance between edge preservation, noise suppression, and resilience to illumination changes, thus ensuring reliable detection across varying conditions.

Table 1. Parameter setup with numerical values and optimal selections for apple detection

Parameter	Description	Range Tested	Best Value
α	Sigmoid sharpness for smoothing	5, 10, 15, 20, 25	15
T	Intensity threshold in sigmoid	0.05, 0.1, 0.15, 0.2	0.1
$|\Omega|$	Local window size for smoothing	$3 \times 3, 5 \times 5, 7 \times 7$	5×5
ϵ	Constant to avoid $\log(0)$	$1e^{-10}, 1e^{-8}, 1e^{-6}$	1×10^{-8}
β	Entropy weight exponent	1.0, 1.2, 1.5, 1.8, 2.0	1.5
τ_E	Threshold for edge map	0.2, 0.3, 0.4, 0.5, 0.6	0.4
r	Radius for local entropy window	1, 2, 3	2

Figure 2 presents representative apple images used to evaluate the proposed fuzzy logic-based detection model. The dataset includes diverse visual challenges, such as varied lighting conditions, occlusion by foliage, changes in viewing angles, and complex backgrounds. These variations simulate real-world scenarios in which detection systems have to perform accurately despite environmental variability. Apples were detected using bounding boxes applied to each image processed by the model.

Figure 3 presents the step-by-step workflow of the proposed apple detection model, designed to accurately localize apples in diverse image conditions. The process began with image acquisition to capture raw apple images. This was followed by preprocessing and smoothing to enhance image quality by reducing noise and improving clarity for effective feature extraction. The next step involved color space transformation, converting the image into a more suitable color model to emphasize discriminative features, particularly the red-green hues of apples. Edge detection was then performed using the Gradient Weighted Edge (GWE) method to leverage the gradient information to precisely delineate object boundaries. The final detection and localization stage identifies the segmented apple region, hence demonstrating the capacity of the model for accurate fruit isolation. Collectively, this pipeline integrates image enhancement, chromatic feature extraction, and edge-based segmentation to ensure reliable apple detection.

Figure 4 demonstrates the effectiveness of the model by means of the visual comparison of the originals of the input images printed on the left and the obtained results of the detection in the right column. The initial pictures captured a diverse set of real-world situations such as different lightings, partial road fades caused by vegetation as

well as complicated backgrounds resulted from manmade orchards. Colored boxes were used as a representation of detected apples in the output images (marked with yellow), as this color showed its maximum difference among the red fruits and green surroundings. The performance of the model was facilitated by an organized pipeline including adaptive preprocessing, color transformation, and GWE-based edge detection. The method increased sensitivity, especially when handling a difficult situation.
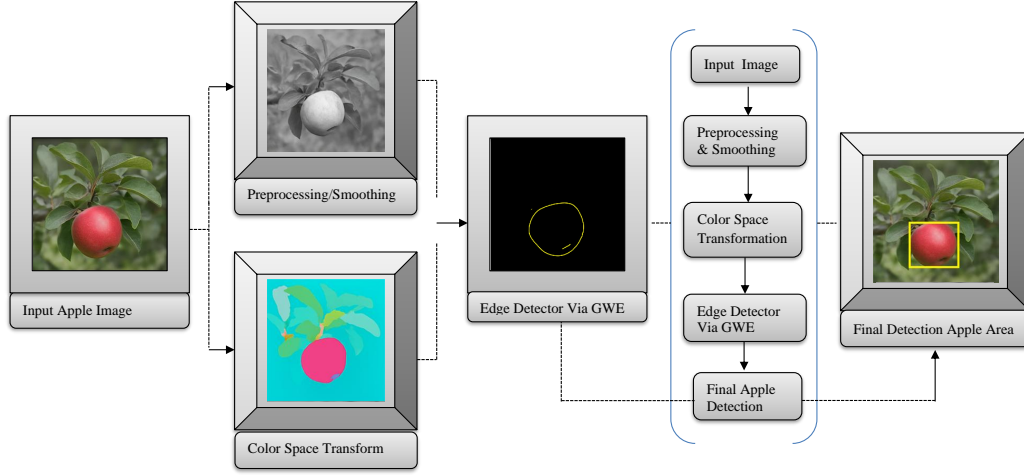

Figure 3. Workflow of the proposed apple detection model, showing the stages of image acquisition, preprocessing, color space transformation, edge detection with the Gradient Weighted Edge (GWE) method, and final apple region localization
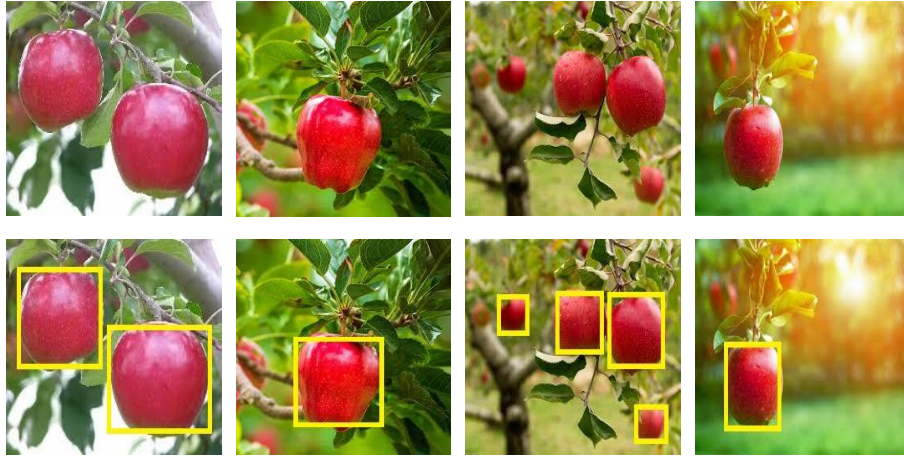

Figure 4. Results of apple detection: The first column displays the original input images whereas the second column presents the output of the proposed model with detected apple regions highlighted by yellow bounding boxes

The model has its key strengths in terms of being able to detect the partially obscured apples, being used in a variety of sized, positioned, and illuminated apples, and generating fewer false positives by excluding non-apple areas. The bounding boxes do not only visually represent reliable instances of detection but also reflect the model confidence. In general, the findings testify to the strength of the system and its real-life usefulness in actual agricultural applications, including yield estimation, automated harvesting, and orchard surveillance, where recognizing targets to a high degree of precision and sensitivity in natural conditions plays a critical role in it. The table below demonstrates that a combination of color analyzing, edge-sensitive processing, and adaptive thresholding could equip the model with a high degree of detection capability.

4.1 Statistical Results of the Proposed Apple Detection Model

Table 2 presents an elaborated statistical evaluation of the proposed apple detection model to confirm both robustness and reliability. The model achieved a Precision of 0.97, indicating highly accurate identification of apples

among all detected positives. A Recall of 0.95 demonstrated that most actual apple instances were successfully detected, while the F1-Score of 0.96 confirmed a strong trade-off between precision and recall. The Intersection over Union (IoU) score of 0.91 further validated the spatial consistency between predicted apple regions and ground truth annotations.

Table 2. Statistical results of the proposed apple detection model

Metric	Our Model (Best Value)	Interpretation
Precision (P)	0.97	High accuracy of apple detection
Recall (R)	0.95	Most actual apples detected
F1-Score	0.96	Strong balance between P and R
IoU	0.91	High pixel-wise overlap accuracy
MOS	4.8 / 5	Expert visual quality assessment (10 reviewers, 1-5 scale)
NIQE	2.5	Lower indicates better image naturalness
BRISQUE	19.2	Lower value indicates better quality
PSNR (dB)	33.8	High fidelity w.r.t. ground truth apple masks
SSIM	0.96	Strong similarity to ground truth annotations

Table 2 presents an elaborated statistical evaluation of the proposed apple detection model to confirm both robustness and reliability. The model achieved a Precision of 0.97, indicating highly accurate identification of apples among all detected positives. A Recall of 0.95 demonstrated that most actual apple instances were successfully detected, while the F1-Score of 0.96 confirmed a strong trade-off between precision and recall. The Intersection over Union (IoU) score of 0.91 further validated the spatial consistency between predicted apple regions and ground truth annotations.

As regards subjective evaluation, the Mean Opinion Score (MOS) was obtained through a controlled experiment involving 10 independent human evaluators with expertise in image analysis. Each evaluator rated the visual quality of segmented apple regions on a 5-point scale (1 = poor, 5 = excellent). The aggregated MOS value of 4.8/5 indicates excellent perceptual quality and strong consensus among reviewers.

Regarding objective metrics involving peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM), the reference images were the manually annotated ground truth apple masks provided in the dataset. The PSNR value of 33.8 decibel (dB) indicates high-fidelity segmentation output with minimal noise relative to the reference whereas the SSIM score of 0.96 demonstrates strong structural similarity to ground truth annotations.

Besides, perceptual quality metrics, like natural image quality evaluator (NIQE) with a score of 2.5 and blind/referenceless image spatial quality evaluator (BRISQUE) with a score of 19.2, confirm low levels of distortion and high naturalness of the detection results. The evaluation process therefore combines both subjective (MOS) and objective (PSNR, SSIM, NIQE, and BRISQUE) measures to ensure comprehensive and transparent performance assessments. Visual evidence of intermediate processing steps is provided in Figure 3 to illustrate the contributions of each stage of the pipeline.

In summary, these analysis results highlight that the proposed model is accurate, perceptually reliable, and suitable for practical orchard environments.

5 Conclusions

Through synergetic integration of the fuzzy logic theory into the self-developed or molded mathematical set, in this study, a powerful and effective framework for detecting apples is proposed to fit the realistic agricultural imagery. In contrast to the methodology of other standard approaches, here adaptive smoothing is combined with entropy-weighted edge detector and directional tracing contours; therefore, methods adopted subsequently could lead to higher resistance levels to common difficulties such as noise, highly lit and darkened environments, and complicated natural backgrounds. The fuzzy logic module allows simple and soft classification as well as uncertainty modelling; this improves the reliability of detection in situations of ambiguity. In addition, the entropy-oriented mechanisms will drive the process of segmentation to more informative areas and directional contour tracking will allow exact localization of the object boundaries. The effectiveness and perceptual quality of the model are further confirmed by copious experimental validation against metrics such as Precision, Recall, IoU, PSNR, SSIM, MOS, NIQE, and BRISQUE measured at both objective and subjective levels. This model has a high potential of applications in precision agriculture, automation of farm harvesting, and real-time fruit observation systems. Real-time implementation and multifruit environment application will be carried out in the future.

Although the proposed fuzzy-based apple detection model has positive outcomes, limitations have to be resolved prior to progressing towards future development initiatives. The first weakness is that the model would be sensitive to severe lighting conditions, in other words, high shadow half-way light glare or over-saturation, with concomitant

effects of weakening the precision of edge locations and contour findings. The other limitation is that fuzzy membership functions and thresholds are manually tuned and the model becomes neither flexible nor suitable in various data sets and in varied environments. Despite these limitations, future work will explicitly focus on developing illumination-invariant preprocessing strategies. For instance, integrating Retinex-based enhancement, histogram equalization, or adaptive illumination normalization could significantly improve edge precision and contour robustness under severe shadows or glares. Instead of relying on manually tuned fuzzy membership functions, optimization-driven techniques such as genetic algorithms or reinforcement learning will be investigated to automate the selection and adjustment of fuzzy rules. This automation will ensure adaptability across different datasets and diverse environmental conditions. Another direction will be the exploration of hybrid neuro-fuzzy systems, where the learning capability of neural networks could be combined with the interpretability of fuzzy logic to dynamically adjust thresholds and rules in real time. Such extensions would not only mitigate the current shortcomings but also extend the general capacity of the model, hence rendering it robust and scalable for deployment in dynamic agricultural settings.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that they have no conflicts of interest.

References

[1] Y. Lin, A. Li, H. Li, Y. Shi, and X. Zhan, "Gpu-optimized image processing and generation based on deep learning and computer vision," *J. Artif. Intell. Gen. Sci.*, vol. 5, no. 1, p. 39–49, 2024. https://doi.org/10.60087/jaigs.v5i1.162

[2] I. Hussain, H. Ali, M. S. Khan, S. Niu, and L. Rada, "Robust region-based active contour models via local statistical similarity and local similarity factor for intensity inhomogeneity and high noise image segmentation," *Inverse. Probl. Imag.*, vol. 16, no. 5, pp. 1113–1136, 2022. https://doi.org/10.3934/ipi.2022014

[3] R. Archana and P. E. Jeevaraj, "Deep learning models for digital image processing: A review." *Artif. Intell. Rev.*, vol. 57, no. 1, p. 11, 2024. https://doi.org/10.1007/s10462-023-10631-z

[4] I. Hussain and J. Muhammad, "Efficient convex region-based segmentation for noising and inhomogeneous patterns," *Inverse. Probl. Imag.*, vol. 17, no. 3, pp. 708–725, 2023. https://doi.org/10.3934/ipi.2022074

[5] M. Shi and I. Hussain, "Improved region-based active contour segmentation through divergence and convolution techniques," *AIMS Math.*, vol. 10, no. 1, pp. 654–671, 2025. https://doi.org/10.3934/math.2025029

[6] L. Zeng, Y. Chen, Y. Yang, and Z. Pan, "Edge-aware image smoothing via weighted sparse gradient reconstruction," *Signal. Image. Video. P.*, vol. 17, no. 8, pp. 4285–4293, 2023. https://doi.org/10.1007/s11760-023-02661-5

[7] R. Matsuoka and M. Okuda, "Beyond staircasing effect: Robust image smoothing via $\ell 0$ gradient minimization and novel gradient constraints," *Signals*, vol. 4, no. 4, pp. 669–686, 2023. https://doi.org/10.3390/signals4040037

[8] S. K. Yadav and K. Sarawadekar, "Robust multi-scale weighting-based edge-smoothing filter for single image dehazing," *Pattern Recogn.*, vol. 149, p. 110137, 2024. https://doi.org/10.1016/j.patcog.2023.110137

[9] Z. Al-Ameen, "Directional variance-based algorithm for digital image smoothing," *Sci. Tech. J. Inf. Technol. Mech. Opt.*, vol. 25, no. 1, pp. 78–86, 2025. https://doi.org/10.17586/2226-1494-2025-25-1-78-86

[10] Y. Zhang and C. Wu, "Unsupervised camouflaged object segmentation as domain adaptation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, Paris, France, 2023, pp. 4336–4346. https://doi.org/10.1109/ICCVW60793.2023.00468

[11] B. Ez-zahouani, A. Teodoro, O. El Kharki, L. Jianhua, I. Kotaridis, X. Yuan, and L. Ma, "Remote sensing imagery segmentation in object-based analysis: A review of methods, optimization, and quality evaluation over the past 20 years," *Remote Sens. Appl: Soc. Environ.*, vol. 32, p. 101031, 2023. https://doi.org/10.1016/j.rsase.2023.101031

[12] C. Bai, X. Bai, and K. Wu, "A review: Remote sensing image object detection algorithm based on deep learning," *Electronics*, vol. 12, no. 24, p. 4902, 2023. https://doi.org/10.3390/electronics12244902

[13] H. Yi, B. Liu, B. Zhao, and E. Liu, "Small object detection algorithm based on improved YOLOv8 for remote sensing," *IEEE J-STARS*, vol. 17, pp. 1734–1747, 2024. https://doi.org/10.1109/JSTARS.2023.3339235

[14] J. Zhang, J. Xie, F. Zhang, J. Gao, C. Yang, C. Song, W. Rao, and Y. Zhang, "Greenhouse tomato detection and pose classification algorithm based on improved YOLOv5," *Comput. Electron. Agr.*, vol. 216, p. 108519, 2024. https://doi.org/10.1016/j.compag.2023.108519

[15] H. Wang, J. Feng, and H. Yin, "Improved method for apple fruit target detection based on YOLOv5s," *Agriculture*, vol. 13, no. 11, p. 2167, 2023. https://doi.org/10.3390/agriculture13112167

[16] S. Kılıçarslan, E. Dönmez, and S. Kılıçarslan, "Identification of apple varieties using hybrid transfer learning and multi-level feature extraction," *Eur. Food Res. Technol.*, vol. 250, no. 3, pp. 895–909, 2024. https://doi.org/10.1007/s00217-023-04436-1

[17] X. Cao, H. Gao, T. Qin, M. Zhu, P. Zhang, and P. Xu, "Boundary aware microscopic hyperspectral pathology image segmentation network guided by information entropy weight," *Front. Oncol.*, vol. 15, p. 1549544, 2025. https://doi.org/10.3389/fonc.2025.1549544

[18] H. M. Zangana, A. K. Mohammed, and F. M. Mustafa, "Advancements in edge detection techniques for image enhancement: A comprehensive review," *Int. J. Artif. Intell. Robot.*, vol. 6, no. 1, pp. 29–39, 2024. https://doi.org/10.25139/ijair.v6i1.8217

[19] Y. Gupta, R. Heintzmann, C. Costa, R. Jesus, and E. Pinho, "Deep learning-enhanced automated mitochondrial segmentation in FIB-SEM images using an entropy-weighted ensemble approach," *PloS ONE*, vol. 19, no. 11, p. e0313000, 2024. https://doi.org/10.1371/journal.pone.0313000

[20] H. S. Gill and B. S. Khehra, "Apple image segmentation using teacher learner based optimization based minimum cross entropy thresholding," *Multimed. Tools Appl.*, vol. 81, no. 8, pp. 11 005–11 026, 2022. https://doi.org/10.1007/s11042-022-12093-x

[21] Y. Wang, N. Li, J. Zhang, Y. Ma, S. Sun, and L. Kong, "Multi feature apple grading based on entropy method and MLP," in *2024 IEEE 4th International Conference on Power, Electronics and Computer Applications (ICPECA)*, Shenyang, China, 2024, pp. 1136–1143. https://doi.org/10.1109/ICPECA60615.2024.10471090