

Acta Technologica Agriculturae 3
Nitra, Slovaca Universitas Agriculturae Nitriae, 2026, pp. 167–178

COMPARATIVE STUDY OF RED-GREEN-BLUE (RGB), HUE-SATURATION-VALUE (HSV), CIE L*A*B* (LAB), AND LUMINANCE-CHROMINANCE (YCBCR) COLOUR SPACES FOR FRUIT SEGMENTATION IN AGRICULTURAL IMAGES

Maxwell SALAZAR^{*1}, Paola PORTERO², Santiago PEREZ²

¹Instituto Superior Tecnológico Rumiñahui, Electronic Department, Sangolquí, Ecuador

²Instituto Superior Tecnológico Rumiñahui, Electricity Department, Sangolquí, Ecuador,
paola.portero@ister.edu.ec (P.P.), santiago.perez@ister.edu.ec (S.P.)

*correspondence: maxwell.salazar@ister.edu.ec

Fruit detection and segmentation are key components of agricultural automation pipelines, enabling tasks such as yield estimation, quality sorting, and robotic harvesting in resource-constrained environments. This study evaluates four colour spaces – RGB, HSV, CIE L*a*b* (LAB), and YCbCr – for threshold-based fruit segmentation in precision agriculture and robotic harvesting. It specifically investigates which colour space maintains the highest accuracy and stability under variable indoor illumination. A dataset of 144 images featuring fruits from four colour categories (red, green, orange, and yellow) was utilised. Colour thresholds were calibrated through manual region-of-interest selection and applied using chromatic thresholding and morphological refinement, resulting in 2304 mask evaluations based on intersection over union (IoU), precision, recall, and F1-score. Clear and statistically significant differences were found among colour spaces: HSV achieved the highest overall performance (F1-score = 0.972, IoU = 0.947), followed by RGB (F1-score = 0.823), LAB (F1-score = 0.714), and YCbCr (F1-score = 0.682), confirmed by non-parametric statistical testing. HSV is established as the most accurate and consistent colour representation for fruit segmentation under the proposed experimental protocol, providing practical guidance for lightweight, real-time segmentation pipelines in agricultural automation systems operating under variable illumination conditions.

Keywords: fruit segmentation; colour spaces; agricultural computer vision; chromatic thresholding

Agriculture is undergoing a rapid technological transformation through automation, robotics, and artificial intelligence, enabling intelligent systems that improve productivity and reduce labour costs in crop production (Alaudeen et al., 2024; An et al., 2022). Robotic harvesting platforms require reliable perception systems to detect fruits accurately before performing picking operations (Afsah-Hejri et al., 2022; Bac et al., 2014), yet fruit detection in natural environments remains challenging due to variable illumination, occlusions, and natural variability in fruit appearance (Li et al., 2022; Lv et al., 2022). Computer vision has emerged as one of the most effective technologies to address these challenges (Cubero et al., 2011; Sa et al., 2016), with machine learning and deep learning approaches significantly improving detection performance (Hamuda et al., 2017; Toyman and Kuscu, 2014; Mureşan and Oltean, 2018); however, they often demand substantial computational resources unfeasible for embedded processors in agricultural robots (Kirillov et al., 2023; Gao et al., 2020). In this context, traditional colour-based segmentation techniques remain highly relevant due to their simplicity, low computational cost, and suitability for real-time implementation on resource-constrained hardware (Hamuda et al., 2016; Riehle et al., 2020). Recent work has demonstrated that even object detection architectures based on deep learning – such as YOLO – are being adapted for real-time fruit detection in harvesting robots (Aljaafreh et al., 2023), underscoring the growing demand for computationally efficient vision

pipelines in agricultural automation. Nevertheless, classical colour-based thresholding remains a competitive baseline when hardware constraints are severe. The effectiveness of colour segmentation depends critically on the colour representation adopted (Gonzalez and Woods, 2018). The standard RGB is sensitive to illumination variations because luminance and chromaticity are coupled within each channel. Alternative spaces such as HSV, LAB, and YCbCr partially decouple luminance from chromatic components, which can improve robustness under variable lighting conditions (Kang and Chen, 2020; Holý et al., 2025). Previous comparative studies have highlighted HSV's superiority for real-time systems (Hamuda et al., 2016; Moreira et al., 2022), LAB's advantages in complex backgrounds (Pujari et al., 2014), and RGB's sensitivity to illumination (Hu et al., 2016). Deep learning models achieve state-of-the-art segmentation performance, but at significant computational cost that constrains their deployment in embedded agricultural systems (Zhang et al., 2024). Despite growing interest in this field, a systematic and statistically validated comparison of colour spaces specifically targeting low-complexity fruit segmentation pipelines suitable for embedded robotic harvesting remains underexplored.

The central research question of this study is which colour space – among RGB, HSV, LAB, and YCbCr – provides the most accurate, stable, and computationally efficient representation for threshold-based fruit segmentation across multiple fruit colours under variable indoor illumination. This study directly

extends the problem of reliable fruit detection in constrained robotic harvesting systems identified by Bac et al. (2014), proposing a reproducible and low-complexity segmentation framework as a practical alternative to deep learning pipelines where hardware constraints are a limiting factor.

The novel contributions of this work are:

- a) a systematic and reproducible comparative evaluation of four colour spaces for threshold-based fruit segmentation using a structured multi-colour dataset;
- b) the application of non-parametric statistical testing (Friedman test with Wilcoxon signed-rank post-hoc analysis and Bonferroni correction) to rigorously validate observed performance differences;
- c) practical guidance for colour space selection in lightweight, real-time agricultural segmentation systems, particularly in scenarios where deep learning is not feasible due to computational constraints.

The remainder of this paper is organised as follows. The Material and Methods section describes the image dataset, colour space definitions, calibration procedure, segmentation pipeline, and evaluation metrics. The Results and Discussion section presents quantitative performance comparisons across colour spaces and fruit colour categories and interprets the findings in the context of related work. The Conclusion section summarises the key insights and directions for future research.

Material and Methods

Dataset

Experiments were conducted on a custom image dataset comprising 144 colour images, assembled specifically for this study to ensure controlled variability in illumination, fruit colour, and background conditions representative of indoor laboratory harvesting scenarios. The dataset was constructed rather than sourced from a public repository because, to the best of the authors' knowledge, no existing public dataset simultaneously covers the four target fruit colour categories (red, green, orange, and yellow) under the variable natural indoor lighting conditions required

to evaluate colour-space robustness for threshold-based segmentation. Images were captured with a 1080P FHD USB webcam (resolution: 1920×1080 pixels) mounted on a fixed support at a distance of approximately 50 cm from the target objects and subsequently resized to 640×480 pixels for processing consistency. All images were acquired in a controlled indoor laboratory environment under variable natural lighting conditions (diffuse daylight through a window, without artificial supplementary illumination), which introduced realistic and representative illumination variability – including shadows and uneven brightness – without simulating outdoor field conditions. Fruits were placed individually and in groups on a plain neutral-coloured surface to minimise background interference while preserving realistic occlusion scenarios.

The dataset contains four fruit colour categories: red, green, orange, and yellow, with 36 images per category (total: 144 images). Each image contains one or more fruits of the target colour category, resulting in a multi-instance segmentation scenario per image. No train/test split was applied in the traditional machine learning sense, as the proposed method is non-parametric and requires no training phase: colour thresholds were calibrated once on a single reference image and subsequently applied without modification to all 144 images. The 144-image corpus therefore constitutes the full evaluation set, and no data partitioning was required.

Ground truth binary masks were generated for each image and each fruit colour category. A ground truth binary mask is a pixel-level annotation image of the same spatial resolution as the input image (640×480 pixels), in which pixels belonging to the target fruit region are assigned a value of 1 (foreground) and all remaining pixels are assigned a value of 0 (background). Ground truth masks were generated for all 144 images across the four fruit colour categories, yielding 576 masks in total ($144 \text{ images} \times 4 \text{ colour categories}$). Since each of the 576 masks was evaluated independently against the predictions of each of the four colour spaces (RGB, HSV, LAB, YCbCr), the total number of mask evaluations is $576 \times 4 = 2304$ evaluations. The annotation procedure is described in detail in the Ground truth masks subsection below. Representative

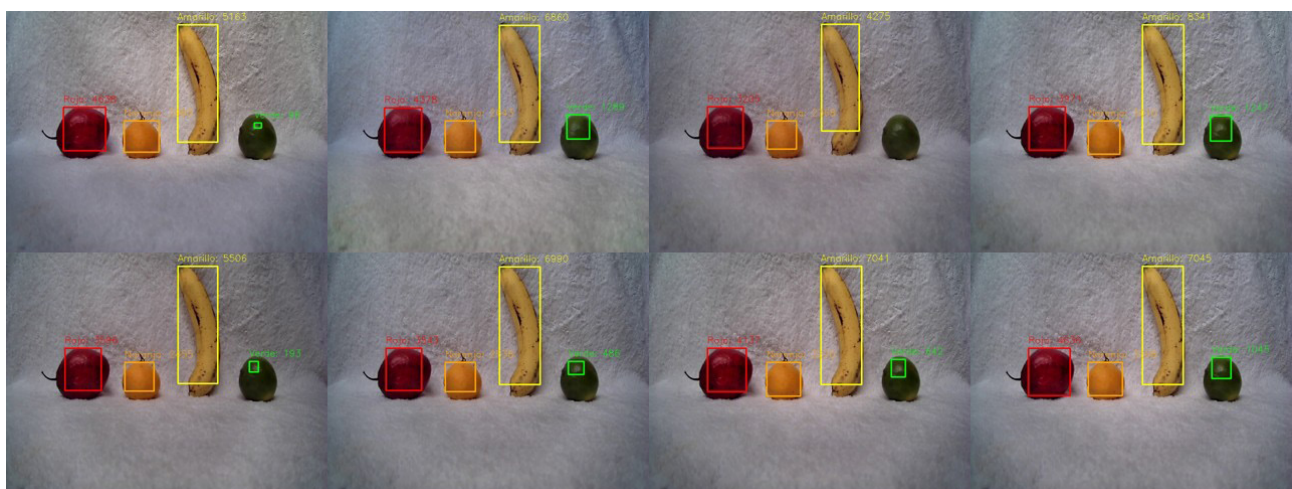


Fig. 1 Representative dataset images showing different fruit colours and illumination conditions

dataset samples illustrating typical illumination conditions, colour variability, and common challenges such as shadows and partial occlusions are shown in Fig. 1.

Colour spaces

Four colour spaces were evaluated in this study: RGB, HSV, CIE L*a*b* (LAB), and YCbCr. Each space represents colour information differently, which directly affects the separability of fruit pixels from background under variable illumination.

RGB (Red, Green, Blue) is the standard representation for digital imaging. Each pixel is described by three additive components (R, G, B) $\in [0, 255]$. Although widely used, RGB is sensitive to illumination changes because luminance and chromaticity are coupled within each channel, making threshold-based segmentation less robust under varying lighting.

HSV (Hue, Saturation, Value) separates chromatic information ($H \in [0^\circ, 360^\circ]$) from saturation ($S \in [0, 1]$) and brightness ($V \in [0, 1]$). This decoupling makes HSV particularly suitable for colour-based segmentation, as fruit colour can be described primarily by the hue channel independently of illumination intensity. HSV thresholding is implemented by defining lower and upper bounds on each of the three channels simultaneously so that a pixel is classified as fruit only when all three channel values fall within their respective calibrated ranges (see Eq. 2).

CIE L*a*b* (LAB) is a perceptually uniform colour space in which L* represents lightness, a* the green–red opponent axis, and b* the blue–yellow opponent axis. LAB is device-independent and designed so that equal numerical differences correspond to approximately equal perceptual differences, which can improve robustness in scenes with complex or variable backgrounds.

YCbCr separates luminance (Y) from blue-difference (Cb) and red-difference (Cr) chrominance components. Originally developed for digital video standards, YCbCr has been applied to skin and colour detection tasks due to its effective decoupling of brightness from colour information.

All colour space conversions were performed using OpenCV 4.8.0 (Bradski, 2000), an open-source computer vision library, via its built-in colour conversion functions (cv2.cvtColor): cv2.COLOR_BGR2HSV, cv2.COLOR_BGR2Lab, and cv2.COLOR_BGR2YCrCb. Since OpenCV loads images in BGR channel order by default, the RGB baseline was applied directly to the loaded BGR image without conversion. For consistency with standard terminology in the literature, this baseline is referred to as “RGB” throughout the manuscript; calibration and thresholding were performed under the same BGR channel ordering.

Range calibration

Colour ranges were calibrated using an interactive region-of-interest (ROI) selection procedure applied to a single representative reference image from the dataset. For each fruit colour category (red, green, orange, and yellow), a rectangular ROI was manually selected to capture pixels representative of that fruit’s colour under typical illumination conditions. The ROI was extracted in each of the four colour representations, and the per-channel bounds were computed from the pixel distribution within the selected region.

Formally, for each channel c in a colour space S , the lower threshold ℓ_c and upper threshold u_c are defined as:

$$\begin{aligned}\ell_c &= \min_{\{p \in \Omega\}} S_c(p), \\ u_c &= \max_{\{p \in \Omega\}} S_c(p)\end{aligned}\quad \text{Eq. (1)}$$

where: p – a pixel, Ω – the set of all pixels inside the selected ROI; $S_c(p)$ – the value of channel c of colour space S at pixel p . This procedure yields, for each fruit colour and each colour space, a pair of threshold vectors $[\ell_1, \ell_2, \ell_3]$ and $[u_1, u_2, u_3]$ defining the chromatic range used during segmentation. The calibrated bounds were stored as a structured configuration file and applied without modification to all 144 evaluation images.

To assess the robustness of this calibration strategy, two approaches were compared:

- min/max bounds derived from a single reference image, as described above;
- 5th–95th percentile bounds computed over a subset of 10 reference images, which reduces the influence of extreme pixel values (outliers) at the ROI boundaries.

The percentile strategy produced marginally lower F1-score for HSV (0.952 vs. 0.954) but improved segmentation stability for orange fruits, where F1-score increased from 0.580 to 0.612 – suggesting that the min/max approach captured some atypical pixels that helped span the full orange colour range. The min/max strategy was retained as the primary calibration method because, under the controlled acquisition conditions of this study, it better captured the full chromatic range of each fruit category without introducing the boundary truncation inherent to percentile-based approaches.

Segmentation pipeline

The segmentation pipeline was implemented in Python using OpenCV 4.8.0 (Bradski, 2000). For each input image, each fruit colour category, and each colour space, the image is first converted into the corresponding colour representation, and a binary mask is computed via chromatic thresholding using the calibrated lower and upper bounds. Specifically, a pixel p is classified as fruit if and only if its channel values fall within the calibrated range in all three channels simultaneously:

$$\ell_c \leq S_c(p) \leq u_c, \forall c \in \{1, 2, 3\} \quad \text{Eq. (2)}$$

This operation was implemented using the cv2.inRange function from OpenCV, which performs per-channel range testing and returns a binary mask of identical spatial dimensions to the input image. Because chromatic thresholding can introduce salt-and-pepper noise (isolated foreground pixels in background regions) and small holes (undetected pixels within fruit regions), the binary masks were refined using morphological operations with a 5×5 square structuring element B . Two sequential operations were applied:

- Morphological opening – erosion followed by dilation – defined as $(M \ominus B) \oplus B$, where $\ominus$ denotes erosion and $\oplus$ denotes dilation. Opening removes small, isolated

foreground regions (noise) while preserving the overall shape of larger connected components.

- Morphological closing – dilation followed by erosion – defined as $(M \oplus B) \ominus B$. Closing fills small holes and gaps within detected fruit regions, producing smoother and more complete region boundaries. Both operations were implemented using cv2.morphologyEx with kernels cv2.MORPH_OPEN and cv2.MORPH_CLOSE, respectively. Figure 2 provides a high-level overview of the complete experimental workflow, from colour range calibration to final mask evaluation.

This diagram outlines the complete workflow of the study, from initial colour range calibration using ROIs to final evaluation metrics, highlighting the sequential steps and data flow across the four colour spaces evaluated.

Ground truth masks

The generation of ground truth binary masks followed a structured two-stage annotation protocol designed to balance annotation efficiency with methodological rigour and reproducibility.

Stage 1 – Automated initialisation via HSV thresholding. To accelerate the annotation process, initial binary masks were generated automatically for each image and each fruit colour category using HSV-based chromatic thresholding. HSV thresholding was selected for initialisation because the hue channel (H) provides a compact and illumination-robust representation of fruit colour, enabling high-recall initial detections that serve as a starting point for manual correction. For each fruit colour category, lower and upper bounds on the H, S, and V channels were defined empirically based on prior knowledge of the target colour ranges (e.g., $H \in [0^\circ, 10^\circ] \cup [170^\circ, 180^\circ]$ for red fruits; $H \in [35^\circ, 85^\circ]$ for green fruits; $H \in [10^\circ, 35^\circ]$ for orange fruits; $H \in [25^\circ, 45^\circ]$ for yellow fruits), and applied using the cv2.inRange function in OpenCV 4.8.0. The resulting masks were then refined morphologically (opening and closing with a 5×5 kernel) to remove isolated noise and fill small gaps before proceeding to manual correction.

Stage 2 – Manual refinement and validation. Each automatically generated mask was subsequently reviewed and corrected manually by a trained annotator using image editing software. Corrections included:

- removing foreground regions that corresponded to background pixels incorrectly detected as fruit (false positives);
- filling foreground gaps within fruit regions that were missed by the automatic thresholding (false negatives).

A pixel region was accepted into the ground truth mask if and only if it unambiguously corresponded to the visible surface of a fruit of the target colour category, without background inclusion. Regions where the fruit boundary was ambiguous due to shadows, overexposure, or

occlusion were conservatively excluded from the annotated foreground.

Inter-annotator agreement. To quantify annotation consistency and support methodological reproducibility, a second independent annotator manually labelled a randomly selected subset of 10 images. Agreement between the two annotators was measured using Cohen's Kappa coefficient (κ), a statistical metric that quantifies the degree of agreement between two raters corrected for the level of agreement expected by chance alone. κ is defined as:

$$\kappa = \frac{(p_o - p_e)}{(1 - p_e)} \quad \text{Eq. (3)}$$

where: p_o – the observed proportion of agreement; p_e – the expected proportion of agreement under chance. The values of κ range from 0 (agreement no better than chance) to 1 (perfect agreement), with values above 0.80 conventionally interpreted as indicating strong agreement (Landis and Koch, 1977). The inter-annotator assessment yielded $\kappa = 0.87$, indicating strong agreement between annotators and supporting the reliability of the annotation protocol.

It should be noted that the use of HSV thresholding to generate initial masks may introduce a systematic bias in favour of HSV-based segmentation, since the ground truth regions were partly shaped by an HSV-based initialisation. This potential limitation is acknowledged and its implications for the interpretation of results are discussed in the Results and Discussion sections.

Evaluation metrics

Predicted binary masks were compared against ground truth masks on a per-pixel basis using four standard segmentation evaluation metrics: intersection over union (IoU), precision, recall, and F1-score. These metrics are defined in terms of three pixel-level counts derived from the comparison between the predicted mask and the ground truth mask for each image and fruit colour category:

- TP (true positives): the number of pixels correctly classified as fruit – i.e., pixels labelled as foreground in both the predicted mask and the ground truth.
- FP (false positives): the number of background pixels incorrectly classified as fruit – i.e., pixels labelled as foreground in the prediction but as background in the ground truth.
- FN (false negatives): the number of fruit pixels missed by the prediction – i.e., pixels labelled as background in the prediction but as foreground in the ground truth.

Intersection over union (IoU), also known as Jaccard index, measures the overlap between the predicted and



Fig. 2 Experimental pipeline overview. Calibration → segmentation in multiple colour spaces → morphological refinement → evaluation against ground truth

ground truth fruit regions relative to their union. It is the most stringent of the four metrics, penalising both over-segmentation and under-segmentation simultaneously:

$$IoU = \frac{TP}{(TP + FP + FN)} \quad \text{Eq. (4)}$$

Precision measures the proportion of predicted fruit pixels that are truly fruit, quantifying the rate of false positive detections:

$$precision = \frac{TP}{(TP + FP)} \quad \text{Eq. (5)}$$

Recall measures the proportion of actual fruit pixels that were correctly detected, quantifying the rate of missed detections (false negatives):

$$recall = \frac{TP}{(TP + FN)} \quad \text{Eq. (6)}$$

F1-score is the harmonic mean of precision and recall, providing a single balanced measure that penalises both over- and under-segmentation. It is equivalent to the Dice similarity coefficient when applied to binary segmentation masks:

$$F1 = \frac{2 \times precision \times recall}{(precision + recall)} = \frac{2TP}{2 \times TP + FP + FN} \quad \text{Eq. (7)}$$

To avoid numerical instability in degenerate cases where the denominator equals zero (i.e., empty predicted and ground truth masks), a small additive constant $\varepsilon = 10^{-6}$ was incorporated into all denominators during implementation. Metrics were computed independently for each of the 2304 mask evaluations (144 images $\times$ 4 fruit colour categories $\times$ 4 colour spaces) and subsequently aggregated as mean values per colour space, per fruit colour category, and as a fruit colour $\times$ colour space interaction matrix of mean F1-scores.

Experimental protocol

The complete experimental procedure consists of four sequential stages:

- Calibration: colour bounds are calibrated once on a single reference image using manual ROI selection, producing a structured configuration file with the lower and upper threshold vectors for each fruit colour category and each colour space.
- Segmentation: the calibrated bounds are applied without modification to every image in the evaluation dataset, generating one predicted binary mask per image, fruit colour category, and colour space – totalling 2304 predicted masks.
- Morphological refinement: each predicted mask is post-processed using sequential morphological opening and closing operations with a 5×5 structuring element, as described in the Segmentation pipeline subsection.
- Evaluation: refined predicted masks are compared against ground truth masks using IoU, precision, recall, and F1-score, yielding 2304 individual metric

evaluations. Results are aggregated and reported as global mean metrics per colour space and per fruit colour category, complemented by qualitative visualisations of representative segmentation outputs.

The full pipeline was implemented in Python 3.12, using OpenCV 4.8.0 for image processing, NumPy for numerical operations, pandas for tabular data management, and matplotlib for result visualisation, executed on a Windows 10 operating system.

Statistical analysis

To determine whether the observed performance differences among the four colour spaces were statistically significant, a non-parametric inferential analysis was conducted on the per-image F1-scores. Non-parametric tests were selected because the distribution of F1-scores across images cannot be assumed to follow a normal distribution, particularly given the multimodal behaviour observed across fruit colour categories.

The Friedman test was applied as the primary global test. The Friedman test is a non-parametric equivalent of the repeated-measures analysis of variance (ANOVA), designed for comparing three or more related groups – in this case, the four colour spaces evaluated on the same set of images. It ranks the performance values within each observation (image) across groups and tests the null hypothesis that all groups perform equivalently. A statistically significant Friedman test indicates that at least one colour space differs from the others in terms of F1-score distribution, without specifying which pairs differ.

Following a significant Friedman test result, post-hoc pairwise comparisons were conducted using the Wilcoxon signed-rank test, a non-parametric test for paired observations that evaluates whether the median difference in F1-score between two colour spaces is significantly different from zero. Since six pairwise comparisons were performed simultaneously (one for each pair among four colour spaces), a Bonferroni correction was applied to control the family-wise error rate. Under this correction, the adjusted significance threshold for each individual comparison is $\alpha = 0.05/6 \approx 0.0083$. All reported p -values are Bonferroni-corrected, and comparisons with corrected $p < 0.0083$ are considered statistically significant.

Results and Discussion

Global performance by colour space

Table 1 summarises the global average segmentation performance metrics for each colour space, aggregated across all 144 images and the four fruit colour categories (2304 mask evaluations). HSV achieved the highest performance across all four metrics, with a mean F1-score of 0.972 and a mean IoU of 0.947, substantially outperforming the remaining colour spaces. RGB ranked second with a mean F1-score of 0.823, followed by LAB (F1-score = 0.714) and YCbCr (F1-score = 0.682).

Inspection of the individual metrics reveals characteristic trade-offs among colour spaces. HSV is the only method achieving simultaneously high precision (0.986) and high recall (0.960), indicating balanced detection with few false

positives and few missed fruit pixels. RGB achieved near-perfect recall (0.993) at the cost of substantially lower precision (0.774), confirming a systematic tendency towards over-segmentation: the method detects most fruit pixels but also incorrectly labels a significant proportion of background pixels as fruit. LAB and YCbCr both achieved high precision (0.939 and 0.953, respectively) but low recall (0.604 and 0.555, respectively), indicating the opposite behaviour – these spaces are conservative in their detections, missing large portions of fruit pixels while maintaining low false positive rates. The low recall of LAB and YCbCr is attributed to the sensitivity of their chromatic channels to the specific illumination conditions of the dataset: the a^* and b^* axes of LAB and the Cb and Cr components of YCbCr are perceptually defined to decouple luminance from colour, but the fixed min/max thresholds calibrated from a single ROI were insufficient to span the full chromatic variability across the 144 evaluation images, particularly under shadows and uneven illumination. As a result, fruit pixels with atypical brightness or saturation were systematically excluded from the predicted masks, producing low recall values without affecting precision.

Performance by fruit colour

Table 2 summarises the average segmentation performance per fruit colour category, aggregated across all colour spaces and images. Red fruits achieved the highest mean F1-score (0.956), followed by green (0.860) and yellow (0.784). Orange fruits were the most challenging, yielding the lowest mean F1-score (0.591) and the lowest mean IoU (0.456). The relatively strong performance on red and green fruits is attributed to their high chromatic contrast with the neutral background used in the dataset, which facilitates threshold-based separation across multiple colour spaces. Orange fruits present greater difficulty due to tonal similarity with warm-coloured backgrounds and higher chromatic variability across images, which reduces the effectiveness of the fixed threshold strategy.

Table 1 Average segmentation performance per colour space (mean over all images and fruit colours)

Colour space	IoU	Precision	Recall	F1
HSV	0.9472	0.9857	0.9597	0.9719
LAB	0.5815	0.9394	0.6039	0.7145
RGB	0.7722	0.7743	0.9926	0.8228
YCbCr	0.5446	0.9534	0.5549	0.6820

Table 2 Average segmentation performance per fruit colour (mean over all images and colour spaces)

Fruit colour	IoU	Precision	Recall	F1
Yellow	0.6854	0.8891	0.7424	0.7844
Orange	0.4563	0.7789	0.6527	0.5911
Red	0.9203	0.9881	0.9299	0.9559
Green	0.7834	0.9967	0.7860	0.8598

Table 3 Mean F1-score by fruit colour and colour space

Fruit colour	HSV	LAB	RGB	YCbCr
Yellow	0.9763	0.6967	0.9392	0.5253
Orange	0.9193	0.5154	0.3638	0.5659
Red	0.9953	0.9223	0.9924	0.9136
Green	0.9967	0.7236	0.9959	0.7232

F1-score matrix: fruit colour vs. colour space

Table 3 presents the mean F1-score for each combination of fruit colour category and colour space, providing a detailed view of per-category performance across methods. HSV consistently achieved the highest F1-score across all four fruit colour categories, with values of 0.997 (green), 0.995 (red), 0.976 (yellow), and 0.919 (orange). The strong performance of HSV on red and green fruits confirms the advantage of hue-based thresholding for high-contrast chromatic categories. The lower HSV performance on orange fruits (F1-score = 0.919) is consistent with the known spectral overlap between orange hues and warm-toned backgrounds, which reduces the discriminability of the hue channel alone.

RGB performed comparably to HSV on red (0.992) and green (0.996) fruits, but degraded substantially on orange (0.364), reflecting the coupling of luminance and chromaticity that makes RGB thresholding highly sensitive to illumination and background colour. LAB and YCbCr showed moderate performance on red and green fruits (F1-scores in the range

0.72–0.92) but consistently low values on orange and yellow categories, where chromatic variability across images exceeded the fixed threshold ranges. These results indicate that the relative ranking of colour spaces is not uniform across fruit colour categories and that method selection should account for the specific chromatic characteristics of the target fruit.

Qualitative results

Qualitative inspection complements the quantitative metrics by illustrating characteristic segmentation behaviours and failure modes across colour spaces. Figure 3 displays the binary masks and segmented outputs produced by each colour space on a representative image, enabling visual comparison of accuracy and robustness to background elements. HSV produced the most complete and precise fruit regions, with well-defined boundaries and minimal false positives. RGB generated masks with high coverage but frequent inclusion of background pixels in bright or warm-toned regions, consistent with its high recall and low precision. LAB and YCbCr produced conservative masks, frequently missing fruit pixels

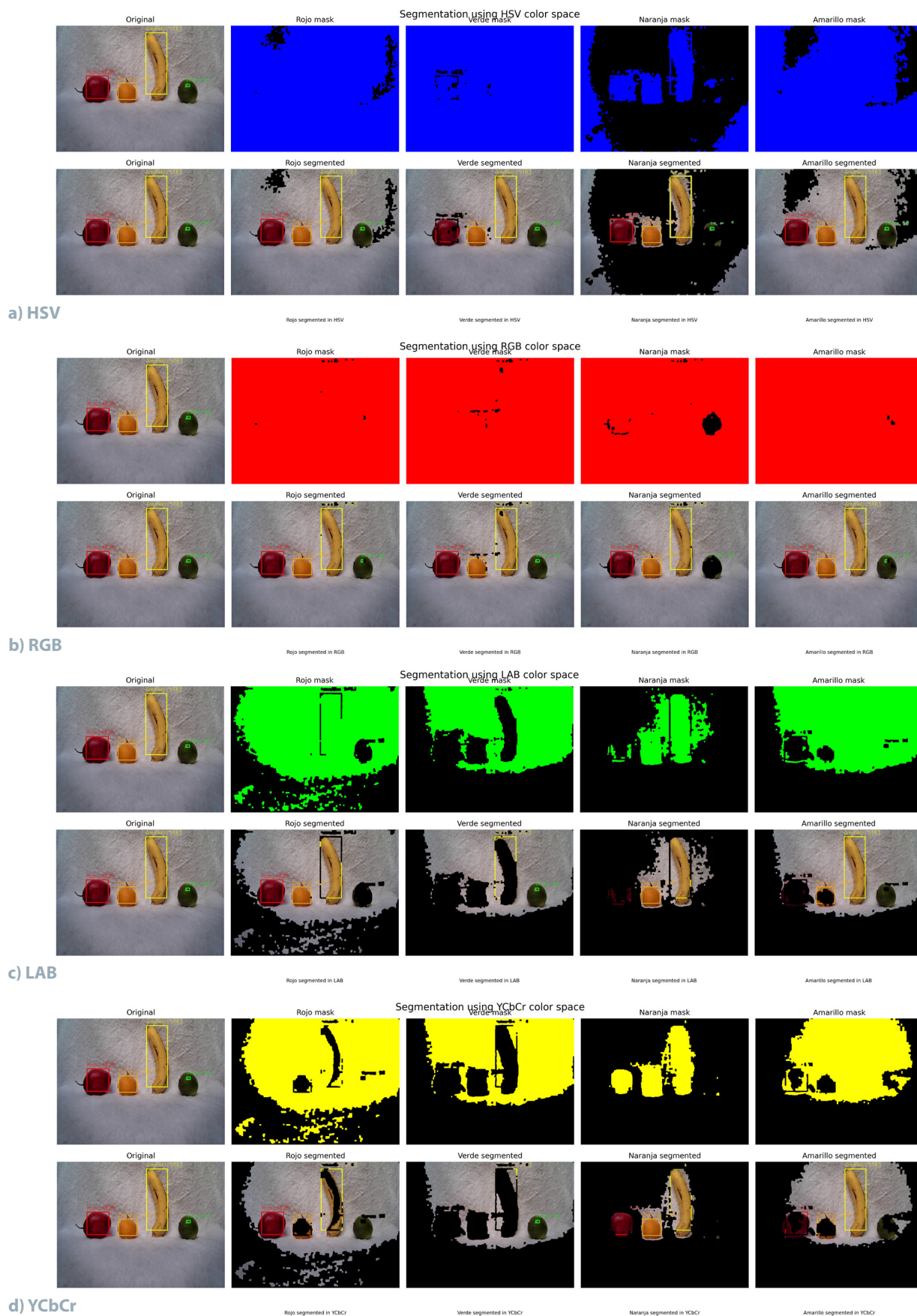


Fig. 3 Qualitative segmentation examples across all four colour spaces. Each panel displays masks (top row) and segmented output (bottom row) for the reference image

in shadowed or chromatically variable regions, consistent with their low recall values. These qualitative patterns reinforce the quantitative findings and highlight that all four colour spaces are evaluated on static images acquired under controlled indoor conditions, with a fixed resolution of 640×480 pixels. The potential effect of resolution changes on segmentation performance was not evaluated in this study; however, since threshold-based segmentation operates on per-pixel chromatic values independently of spatial resolution, the impact of moderate resolution changes is expected to be limited to boundary precision rather than overall chromatic separability.

Statistical analysis

To assess whether the observed performance differences among colour spaces are statistically significant under matched conditions, a Friedman test was applied to the F1-scores using matched blocks defined by (image, fruit colour), treating the four colour spaces as repeated-measures conditions. The significance level was set at $\alpha = 0.05$. The Friedman test yielded a test statistic of $\chi^2 = 312.19$ with 3 degrees of freedom ($df = 3$), corresponding to a critical value of $\chi^2_{\text{crit}}(3, 0.05) = 7.815$. Since the observed statistic far exceeds the critical value, the null hypothesis of equal distributions across colour spaces was rejected ($p = 2.29 \times 10^{-67}$; Kendall's $W = 0.631$), confirming that at least one colour space performs significantly differently from the others.

Post-hoc pairwise comparisons were conducted using two-sided Wilcoxon signed-rank tests comparing HSV against each alternative colour space. The Bonferroni-corrected significance threshold for the six simultaneous pairwise comparisons is $\alpha_{\text{corr}} = 0.05/6 \approx 0.0083$ (critical z -value ≈ 2.638 for a two-sided test). All pairwise comparisons yielded corrected p -values substantially below this threshold: HSV vs. RGB ($p_{\text{corr}} = 1.23 \times 10^{-15}$), HSV vs. LAB ($p_{\text{corr}} = 2.40 \times 10^{-28}$), and HSV vs. YCbCr ($p_{\text{corr}} = 2.40 \times 10^{-28}$), confirming that HSV significantly outperforms all three alternative colour spaces under the experimental conditions of this study.

Descriptive analysis and visualization

Figures 4, 5, and 6 visualise the mean IoU, precision/recall, and F1-score per colour space, respectively. HSV achieves the highest IoU (0.947), exceeding RGB (0.772) by 0.175 points and outperforming LAB (0.581) and YCbCr (0.545) by wider margins, confirming that HSV produces masks substantially more aligned with ground truth boundaries. The precision-recall analysis (Fig. 5) reveals a clear trade-off: RGB achieves near-perfect recall (0.993) at the cost of low precision (0.774), confirming systematic over-segmentation, whereas HSV is the only method achieving high values on both axes simultaneously (precision = 0.986, recall = 0.960). The F1-score distribution (Fig. 6) reinforces this ranking: the gap between HSV (0.972)

and RGB (0.823) is 0.149 points – larger than the gap between RGB and the lowest-performing method YCbCr (0.682) – suggesting that HSV operates in a qualitatively different performance tier rather than being merely incrementally superior.

Figures 7 and 8 address consistency across images. HSV's IoU distribution is concentrated in the upper range (IQR = 0.05) with few low-value outliers, whereas RGB shows a wider spread (IQR = 0.12) with values falling below 0.20 in the most challenging images. The Dice coefficient distributions (Fig. 8) corroborate this pattern: HSV shows a median exceeding 0.97 with a tight interquartile range, while LAB and YCbCr exhibit median values near 0.70 with high variance, indicating inconsistent performance across image and fruit colour combinations.

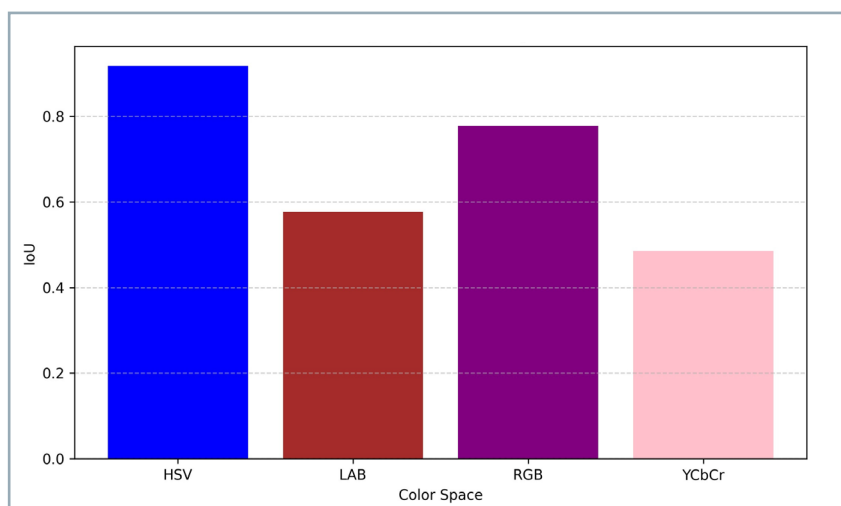


Fig. 4 Average IoU per colour space

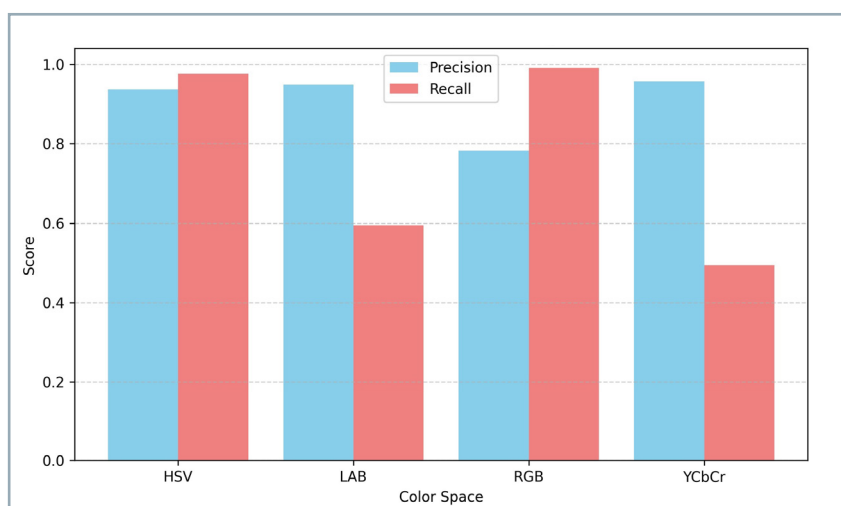


Fig. 5 Average precision and recall per colour space

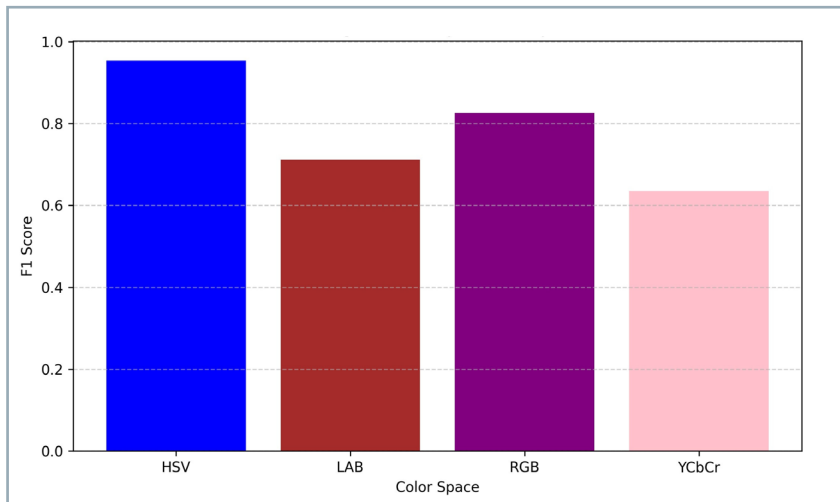


Fig. 6 Average F1-score per colour space

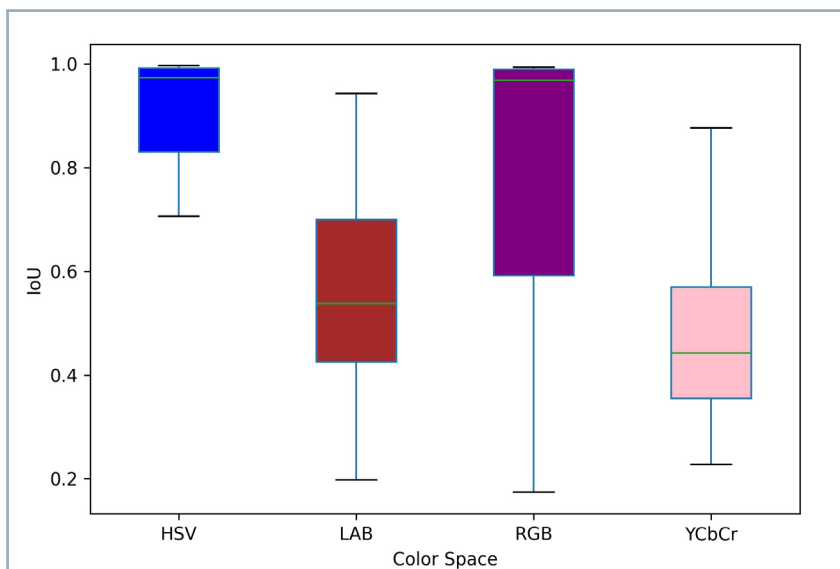


Fig. 7 Distribution of IoU across samples (boxplot)

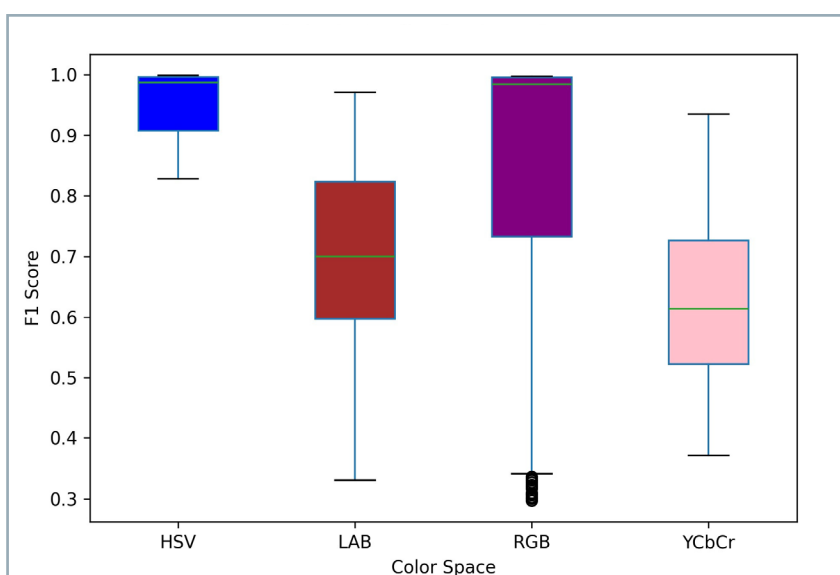


Fig. 8 Distribution of Dice coefficient (F1-score) values across all colour spaces

The experimental results demonstrate clear and statistically significant performance differences among the four evaluated colour spaces for threshold-based fruit segmentation. HSV achieved the highest overall performance (mean F1-score = 0.972, mean IoU = 0.947), attributable to its effective separation of chromatic information (hue channel, H) from luminance (value channel, V), which enables robust colour-based thresholding under variable illumination conditions. This finding is consistent with previous studies reporting HSV's superiority for real-time colour segmentation in agricultural environments (Hamuda et al., 2016; Hu et al., 2016; Moreira et al., 2022). The higher absolute F1-scores observed in the present study relative to some prior works likely reflect the controlled indoor acquisition conditions, which, while realistic in terms of illumination variability, do not replicate the full complexity of outdoor field environments.

RGB ranked second (F1-score = 0.823) but exhibited a characteristic over-segmentation pattern, achieving near-perfect recall (0.993) at the cost of substantially lower precision (0.774). This behaviour is explained by the coupling of luminance and chromaticity within the RGB channels: under variable illumination, background pixels with warm or bright tones can fall within the calibrated RGB threshold range of fruit pixels, producing systematic false positives. LAB performed moderately (F1-score = 0.714), with high precision (0.939) but low recall (0.604), indicating that the fixed min/max thresholds calibrated from a single ROI were insufficient to capture the full chromatic variability of fruit pixels across the 144 evaluation images, particularly under shadow regions where L^* values deviate significantly from the reference. YCbCr showed the lowest overall performance (F1-score = 0.682) despite high precision (0.953), suggesting that its chrominance components (Cb, Cr) do not provide sufficient separability between

fruit and background chromatic distributions under the illumination conditions of this dataset.

Analysis by fruit colour revealed that orange fruits were the most challenging category across all colour spaces (mean F1-score = 0.591), due to spectral overlap between orange hues and warm-toned background elements and higher tonal variability across images. Red and green fruits achieved the highest mean F1-scores (0.956 and 0.860, respectively), consistent with their high chromatic contrast with the neutral background. Qualitative inspection confirmed characteristic failure modes: HSV under-segments in low-illumination regions and struggles with orange hues that overlap background elements; RGB produces false positives in bright and warm-toned backgrounds; LAB misses fruit pixels in high-contrast or shadowed areas; and YCbCr confuses fruits with chromatically similar backgrounds in the Cb–Cr space.

Comparison with machine learning-based approaches

A relevant dimension for contextualising the present results is the comparison of threshold-based colour segmentation with machine learning and deep learning approaches. Convolutional neural network-based detectors – such as YOLO architectures – have demonstrated high accuracy in real-time fruit detection tasks in agricultural settings, achieving F1-scores exceeding 0.90 on challenging outdoor datasets (Aljaafreh et al., 2023; Zhang et al., 2024). However, these methods require annotated training datasets, significant computational resources for training, and hardware accelerators (GPU) for real-time inference on embedded systems. In contrast, the threshold-based HSV pipeline proposed in this study achieved a mean F1-score of 0.972 under controlled conditions without any training phase, requiring only a single reference image for calibration, and is executable on CPU-only embedded hardware with minimal memory footprint. This positions threshold-based HSV segmentation as a competitive baseline for scenarios where computational constraints are severe, training data is unavailable, or rapid deployment without offline training is required. The key limitation of the proposed approach relative to machine learning methods is its sensitivity to illumination changes beyond the calibrated range and its reduced generalisation to unseen backgrounds, fruit varieties, or acquisition conditions – limitations that deep learning models partially overcome through learned feature representations. Future work should include a direct quantitative comparison between the proposed pipeline and lightweight deep learning models (e.g., MobileNet-SSD, YOLOv8-nano) under matched experimental conditions to establish a rigorous performance – complexity trade-off.

Limitations

Several limitations of this study must be acknowledged. First, ground truth masks were generated via HSV-based initialisation followed by manual refinement, which may introduce a systematic annotation bias favouring HSV-detected regions. Although inter-annotator agreement was substantial (Cohen's Kappa = 0.87), a degree of HSV performance overestimation cannot be fully excluded. Second, colour range calibration relied on min/max bounds derived from a single reference image, limiting

robustness to illumination variability beyond the calibrated range; a percentile-based alternative improved stability for orange fruits but slightly reduced overall F1-score, so the min/max method was retained as the primary strategy. Third, no illumination normalisation was applied prior to segmentation; preprocessing techniques such as histogram equalisation or adaptive gamma correction could further improve performance under high-contrast or low-illumination conditions. Fourth, the dataset comprises 144 images captured in a single indoor laboratory environment with a neutral background, which limits the generalisability of the conclusions to more diverse acquisition conditions, backgrounds, and fruit varieties. Evaluation on larger and more varied datasets would strengthen the external validity of the findings. Finally, the RGB baseline was processed in OpenCV's default BGR channel order, which is internally consistent throughout the pipeline but should be noted explicitly for reproducibility purposes.

Conclusion

This study presented a systematic and statistically validated comparative evaluation of four colour spaces – RGB, HSV, CIE L*a*b* (LAB), and YCbCr – for threshold-based fruit segmentation in agricultural images. A dataset of 144 images encompassing four fruit colour categories (red, green, orange, and yellow) was used to assess performance across 2304 mask evaluations using intersection over union, precision, recall, and F1-score as evaluation metrics.

The results clearly establish HSV as the most accurate and consistent colour representation for the proposed segmentation task, achieving a mean F1-score of 0.972 and a mean IoU of 0.947, outperforming RGB (F1-score = 0.823), LAB (F1-score = 0.714), and YCbCr (F1-score = 0.682) by statistically significant margins confirmed through a Friedman test and post-hoc Wilcoxon signed-rank tests with Bonferroni correction. The superiority of HSV is attributed to its effective decoupling of chromatic information from luminance, which provides robust fruit-background separability under variable indoor illumination. RGB exhibited systematic over-segmentation due to luminance–chromaticity coupling, while LAB and YCbCr showed conservative under-segmentation attributed to the limited generalisation of fixed thresholds calibrated from a single reference image.

From a practical standpoint, the proposed pipeline constitutes a reproducible, training-free, and computationally lightweight framework suitable for deployment in resource-constrained embedded robotic harvesting systems where deep learning approaches are infeasible due to hardware limitations. The pipeline requires only a single reference image for colour threshold calibration and operates with minimal memory and processing overhead, making it directly applicable to real-time agricultural automation scenarios. However, it is important to emphasise that the experimental conditions of this study were controlled: images were acquired indoors under variable natural diffuse lighting, on a neutral background, and at a fixed camera distance of 50 cm. The transferability of the reported performance to uncontrolled outdoor field conditions – where illumination intensity, directionality, background diversity, and fruit appearance variability are

substantially greater – cannot be guaranteed without additional validation. Deployment in outdoor environments would require either adaptive threshold recalibration under field conditions or the integration of illumination normalisation preprocessing to maintain the segmentation performance observed in this study.

The three primary scientific contributions of this work are:

1. a reproducible comparative evaluation framework for colour space selection in agricultural fruit segmentation;
2. rigorous non-parametric statistical validation of performance differences across colour spaces and fruit colour categories;
3. practical decision guidance for colour space selection in lightweight real-time segmentation pipelines, including the identification of orange fruits as the most challenging category requiring dedicated attention in future calibration strategies.

Future research directions include:

1. the development of adaptive and robust calibration methods, such as percentile-based threshold estimation over multiple reference images, to improve generalisation to unseen illumination conditions;
2. the integration of illumination normalisation preprocessing – such as histogram equalisation or adaptive gamma correction – prior to segmentation, to extend applicability to outdoor and field environments;
3. the expansion of the evaluation dataset to include additional fruit species, diverse backgrounds, and outdoor acquisition conditions to strengthen the external validity of the conclusions;
4. the integration of shape and texture descriptors as complementary features for hybrid segmentation pipelines that reduce reliance on chromatic thresholding alone;
5. a direct quantitative comparison with lightweight deep learning segmentation models – such as YOLOv8-nano or MobileNet-based architectures – under matched experimental conditions, to establish a rigorous performance – complexity trade-off for embedded agricultural systems.

Acknowledgement

The authors would like to thank the Instituto Tecnológico Superior Rumiñahui for its support and sponsorship of this research. The resources and institutional facilities provided were essential for the successful completion of the experiments presented in this article.

References

- Afsah-Hejri, L., Homayouni, T., Toudeshki, A., Ehsani, R., Ferguson, L., Castro-García, S., 2022. Mechanical harvesting of selected temperate and tropical fruit and nut trees. *Horticultural Reviews* 49, 171–242. <https://doi.org/10.1002/9781119851981.ch4>
- Alaudeen, K.M., Selvarajan, S., Manoharan, H., Jhaveri, R.H., 2024. Intelligent robotics harvesting system process for fruits grasping prediction. *Scientific Reports* 14, 2820. <https://doi.org/10.1038/s41598-024-52743-8>
- Aljaafreh, A., Elzagzoug, E.Y., Abukhait, J., Soliman, A.-H., Alja'afreh, S.S., Sivanathan, A., Hughes, J., 2023. A real-time olive fruit detection for harvesting robot based on YOLO algorithms. *Acta Technologica Agriculturae* 26(3), 121–132. <https://doi.org/10.2478/ata-2023-0017>
- An, Z., Wang, C., Raj, B., Eswaran, S., Raffik, R., Debnath, S., Rahin, S.A., 2022. Application of new technology of intelligent robot plant protection in ecological agriculture. *Journal of Food Quality* 2022, 1257015. <https://doi.org/10.1155/2022/1257015>
- Bac, C.W., van Henten, E.J., Hemming, J., Edan, Y., 2014. Harvesting robots for high-value crops: State-of-the-art review and challenges ahead. *Journal of Field Robotics* 31(6), 888–911. <https://doi.org/10.1002/rob.21525>
- Bradski, G., 2000. The OpenCV library. *Dr. Dobb's Journal* 25(11), 120–125. https://www.proquest.com/docview/202684726/fulltextPDF/74C289D9CFE6423FPQ/1?accountid=32559&source_type=Trade%20Journals
- Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., Blasco, J., 2011. Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and Bioprocess Technology* 4, 487–504. <https://doi.org/10.1007/s11947-010-0411-8>
- Gao, Z., Shao, Y., Xuan, G., Wang, Y., Liu, Y., Han, X., 2020. Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. *Artificial Intelligence in Agriculture* 4, 31–38. <https://doi.org/10.1016/j.aiia.2020.04.003>
- Gonzalez, R.C., Woods, R.E., 2018. *Digital Image Processing*. 4th ed. Pearson Education, New York, USA. ISBN 978-0-13-335672-4. <https://www.cl72.org/090imagePLib/books/Gonzales,Woods-Digital.Image.Processing.4th.Edition.pdf>
- Hamuda, E., Ginley, B.M., Glavin, M., Jones, E., 2017. Automatic crop detection under field conditions using the HSV colour space and morphological operations. *Computers and Electronics in Agriculture* 133, 97–107. <https://doi.org/10.1016/j.compag.2016.11.021>
- Hamuda, E., Glavin, M., Jones, E., 2016. A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture* 125, 184–199. <https://doi.org/10.1016/j.compag.2016.04.024>
- Holý, M., Cviklovič, V., Tóth, L., Vacho, L., Olejár, M., Kósa, P., Baláži, J., Paulovič, S., 2025. Tomato stem recognition based on YOLO. *Acta Technologica Agriculturae* 28(3), 131–140. <https://doi.org/10.2478/ata-2025-0017>
- Hu, M.-H., Dong, Q.-L., Liu, B.-L., Pan, L.-Q., Walshaw, J., 2016. Image segmentation of bananas in a crate using a multiple threshold method. *Journal of Food Process Engineering* 39(5), 427–432. <https://doi.org/10.1111/jfpe.12233>
- Kang, H., Chen, C., 2020. Fruit detection, segmentation and 3D visualisation of environments in apple orchards. *Computers and Electronics in Agriculture* 171, 105302. <https://doi.org/10.1016/j.compag.2020.105302>
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4015–4026. Paris Convention Centre, Paris, France. https://openaccess.thecvf.com/content/ICCV2023/papers/Kirillov_SegmentAnything_ICCV_2023_paper.pdf
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Li, Y., Feng, Q., Li, T., Xie, F., Liu, C., Xiong, Z., 2022. Advance of target visual information acquisition technology for fresh fruit robotic harvesting: a review. *Agronomy* 12(6), 1336. <https://doi.org/10.3390/agronomy12061336>

- Lv, J., Xu, H., Xu, L., Zou, L., Rong, H., Yang, B., Niu, L., Ma, Z., 2022. Recognition of fruits and vegetables with similar-color background in natural environment: a survey. *Journal of Field Robotics* 39(6), 888–904. <https://doi.org/10.1002/rob.22074>
- Moreira, G., Magalhães, S.A., Pinho, T., dos Santos, F.N., Cunha, M., 2022. Benchmark of deep learning and a proposed HSV colour space models for the detection and classification of greenhouse tomato. *Agronomy* 12(2), 356. <https://doi.org/10.3390/agronomy12020356>
- Mureşan, H., Oltean, M., 2018. Fruit recognition from images using deep learning. *Acta Universitatis Sapientiae, Informatica* 10(1), 26–42. <https://doi.org/10.2478/ausi-2018-0002>
- Pujari, J.D., Yakkundimath, R., Byadgi, A.S., 2014. Recognition and classification of produce affected by identically looking powdery mildew disease. *Acta Technologica Agriculturae* 17(2), 29–34. <https://doi.org/10.2478/ata-2014-0007>
- Riehle, D., Reiser, D., Griepentrog, H.W., 2020. Robust index-based semantic plant/background segmentation for RGB- images. *Computers and Electronics in Agriculture* 169, 105201. <https://doi.org/10.1016/j.compag.2019.105201>
- Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. DeepFruits: a fruit detection system using deep neural networks. *Sensors* 16(8), 1222. <https://doi.org/10.3390/s16081222>
- Toylan, H., Kuseu, H., 2014. A real-time apple grading system using multicolor space. *The Scientific World Journal* 2014, 292681. <https://doi.org/10.1155/2014/292681>
- Zhang, J., Kang, N., Qu, Q., Zhou, L., Zhang, H., 2024. Automatic fruit picking technology: a comprehensive review of research advances. *Artificial Intelligence Review* 57, 54. <https://doi.org/10.1007/s10462-023-10674-2>

