Computers and Electronics in Agriculture 237 (2025) 110593

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

A wearable thumb device for fruit firmness estimation with vision-based tactile sensing

Mashood M. Mohsan [a,b,1], Basma B. Hasanen [a,b,1], Taimur Hassan [c], Lakmal Seneviratne [a,b], Irfan Hussain [a,b,*]

[a] Khalifa University Center for Autonomous Robotic Systems (KUCARS), Khalifa University, United Arab Emirates
[b] Department of Mechanical and Nuclear Engineering, Khalifa University, Abu Dhabi, 127788, United Arab Emirates
[c] Department of Electrical, Computer and Biomedical Engineering, Abu Dhabi University, United Arab Emirates

ARTICLE INFO

Keywords:
Vision-based tactile sensing
Deep learning
Wearable
Agricultural device
Firmness

ABSTRACT

Recent advancements in non-destructive technologies have enabled precise firmness measurement for various fruits, including kiwifruit. However, existing methods remain limited by high costs, environmental sensitivity, and field application impracticality. This work introduces a novel wearable device for estimating non-destructive fruit firmness, combining human tactile interaction with vision-based tactile sensing and edge computing. Worn on the thumb, the device leverages embodied intelligence, merging intuitive human touch with the precision of a vision-based tactile sensor. A single-board computer processes tactile images locally, enabling reliable operation even in remote environments. The device employs our proposed deep learning model for real-time firmness predictions from a single palpation, minimizing repetitive handling and reducing fruit bruising. Its ergonomic, symmetrical design supports comfortable use on either hand, enhancing usability. Compact and portable, the device integrates essential components within a housing measuring 40 mm × 25 mm × 72 mm and weighing only 135 g. Validated through non-destructive ripeness assessments on 'Hayward' Kiwifruit, the device demonstrated a strong correlation between tactile images and firmness values when paired with our proposed model, achieving a coefficient of determination (R^2) of 0.89. This study created a dedicated dataset on Kiwi firmness to support model development and validation. Moreover, this work's proposed dataset and source code is available at https://mashood3624.github.io/WearableDevice/.

1. Introduction

Each year, approximately one billion tons of food are wasted globally, intensifying food insecurity and underscoring the urgent need for sustainable practices in the food industry (Voss et al., 2024). Fruit production is particularly critical, as rising consumer demand for high-quality produce increases pressure to minimize losses. Efficient quality assessment plays a key role in addressing this challenge.

Ripeness is a fundamental measure of fruit quality, directly influencing taste, and marketability. Accurate ripeness evaluation ensures fruits are harvested at their optimal maturity, enhancing flavor, nutritional value, and storage potential. It also reduces waste during distribution and increases consumer satisfaction, providing a competitive edge for producers (Mazhar et al., 2016)(Khalifa et al., 2011a).

For fruits such as mangoes, bananas, and apples, ripeness is often determined by visible skin color changes, with computer vision (CV)-based solutions providing non-invasive and effective assessments (Valiente et al., 2021-10; Rehman et al., 2024). In contrast, fruits such as Kiwis, which show little to no color change during ripening, rely on firmness as a more reliable indicator of maturity and readiness for consumption. This is especially crucial for Kiwifruit, valued for its high nutritional content and economic significance. An accurate assessment of firmness is essential to maintain quality, preserve market appeal, and minimize waste (Nazir et al., 2024; Khan et al., 2023). Moreover, Kiwifruit's susceptibility to bruising during handling highlights the importance of gentle and precise evaluation methods to ensure quality (Ahmadi, 2018; Harker and Hallett, 1994).

Although various non-destructive tools have been explored, many are constrained by operator dependence, environmental sensitivity, and limited adaptability across different fruit types (Abbott et al., 1995; Anjali et al., 2024). Vision-based tactile sensing (VBTS) has recently

* Corresponding author.
E-mail addresses: mashood.mohsan@ku.ac.ae (M.M. Mohsan), basma.hasanen@ku.ac.ae (B.B. Hasanen), taimur.hassan@adu.ac.ae (T. Hassan), lakmal.seneviratne@ku.ac.ae (L. Seneviratne), irfan.hussain@ku.ac.ae (I. Hussain).

[1] Authors equally contributed to this work.

https://doi.org/10.1016/j.compag.2025.110593
Received 17 February 2025; Received in revised form 13 May 2025; Accepted 22 May 2025
Available online 24 June 2025

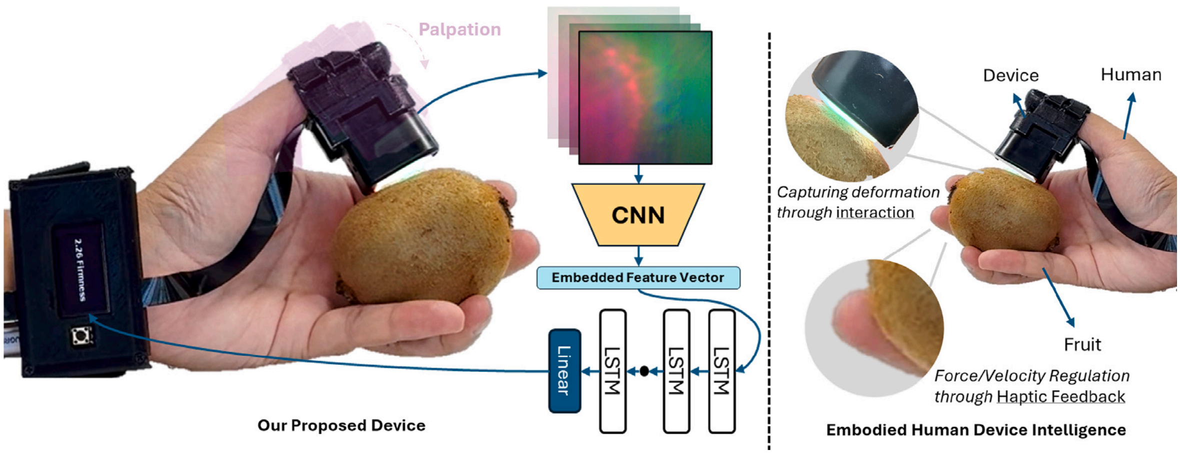

Fig. 1. The proposed wearable device for real-time, and non-destructive fruit firmness estimation. The user palpates a Kiwi, and the proposed model processes the VBTS palpation recording to predict firmness in a non-destructive approach. The right segment illustrates the concept of embodied human-device intelligence, highlighting how haptic feedback and human interaction regulate force during palpation.

emerged as a promising alternative, capturing high-resolution deformation data through soft elastomeric interfaces. When coupled with deep learning, these systems offer improved accuracy and robustness for firmness estimation under variable conditions (Ma et al., 2024; Lin et al., 2023; Yuan et al., 2016, 2017; Mohsan et al., 2025). Building on the strengths of VBTS while addressing the limitations of existing tools, we present a novel wearable device specifically designed for fruits like kiwifruit, where firmness is a key ripeness indicator. The device, worn on the thumb, naturally follows the motion of human touch to apply controlled, gentle pressure to the fruit's surface (Fig. 1). An integrated RGB camera records deformation in the elastomer, and deep learning algorithms process these patterns to predict firmness with high precision. To the best of our knowledge, this is the first system to integrate VBTS principles with human tactile dynamics in a wearable form, enabling non-destructive, intuitive, and accurate firmness assessment.

A key advantage of the device is its efficiency, as it determines Kiwifruit ripeness in a single palpation motion — unlike humans, who often require multiple attempts, increasing the risk of bruising (Rivera et al., 2023; Barrett et al., 1998). The design leverages natural haptic feedback from the supporting fingers, enabling real-time pressure adjustments. In comparison to robotic systems, which often prioritize precision but lack adaptability (Mohsan et al., 2025), the wearable device integrates advanced sensors with the human ability to intuitively adjust force. This hybrid approach ensures accurate, efficient, and gentle handling of delicate fruits like Kiwifruit. By merging automation with human intuition, the device provides a sustainable and reliable solution for enhancing firmness evaluation while minimizing damage, ultimately improving fruit quality.

This paper has the following key contributions:

1. A novel wearable thumb device for non-destructive fruit firmness estimation, integrating a vision-based tactile sensor (VBTS) to capture localized surface deformations during palpation. This AI-driven approach, which combines human intelligence with VBTS analysis, is, to the best of our knowledge, the first of its kind.
2. A deep learning model is proposed and deployed on a compact single-board computer, enabling local video-based inference to estimate firmness by extracting spatial and temporal features from tactile palpation.
3. A "Hayward" Kiwifruit dataset was collected, comprising tactile palpation recordings and penetrometer-based firmness measurements to develop and validate the proposed device. This dataset and the source code of the proposed model will be publicly available upon acceptance.
4. The proposed wearable device was validated for firmness assessment through field experiments, demonstrating its practicality for non-destructive, real-time firmness evaluation in agriculture.

2. Related work

2.1. Traditional methods of fruit firmness estimation

Mechanical devices for evaluating fruit firmness typically utilize compression, puncture, and impact tests (Li et al., 2016). Among invasive methods, the Magness-Taylor (MT) penetrometer remains a widely adopted tool for measuring rupture force by inserting a probe into the fruit (Abbott, 1999). However, the reliance on operator skill in using these devices introduces variability in results (Harker et al., 1996). To address this challenge, advancements such as force gauges mounted on controlled stands have been developed to enhance precision (Jantra et al., 2018). Despite these improvements, the inherently invasive nature of these devices poses a significant limitation, as the tested samples are unusable afterwards.

In contrast, non-invasive mechanical devices, like durometers, assess parameters such as resistance or bioyield force with minimal damage to the fruit. However, their accuracy depends on user technique, and they often require reconfiguration for different fruit varieties, reducing their versatility (Harker et al., 1996). Beyond mechanical approaches, acoustic and vibrational methods have emerged as promising non-invasive alternatives for assessing firmness. Acoustic devices operate by generating sound waves through impact excitation and analyzing the resulting signals to determine fruit firmness (Khalifa et al., 2011b). Vibrational methods, on the other hand, involve generating vibrations and detecting the response, which is influenced by the resonance frequency of the fruit and correlates closely with its firmness. While promising, these methods remain susceptible to environmental factors such as temperature, humidity, and background noise, affecting their reliability in field conditions (Abbott et al., 1995).

Additionally, optical methods provide advanced non-invasive techniques for firmness evaluation, utilizing visible and near-infrared (NIR) spectra to measure various quality attributes. Reflectance-based optical devices capture diffusing reflectance spectra to construct predictive models for firmness-related parameters, as demonstrated by handheld NIR analyzers and portable Vis/NIR spectrometers (Cirilli et al., 2016; Huang et al., 2018). Transmittance-based optical devices complement these approaches by measuring light that passes through the fruit, offering valuable insights into internal quality attributes. However, these systems can be costly and sensitive to variations in fruit surface properties such as color, texture, and shape. Additionally, environmental factors like dust, moisture, and surface damage can introduce inconsistencies, necessitating careful calibration (Anjali et al., 2024).

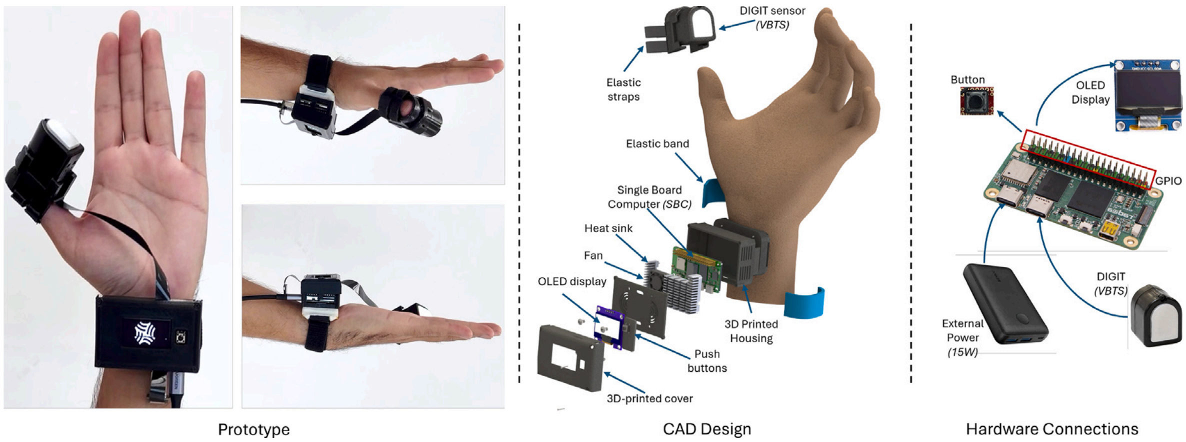

Fig. 2. Development of proposed device (Left) The prototype of proposed device. (Center) A CAD exploded view of our device, illustrating its two main modules: the thumb module with a VBTS at the tip of the thumb and the wrist module containing the main control unit and user interaction interface in the 3D-printed housing. Adjustable elastic straps are used in both modules for secure and comfortable use. (Right) Hardware setup and connections: The Radxa Zero 2 Pro is connected with one button, an OLED display, a DIGIT (VBTS) sensor, and an external power bank of 15 W is used.

2.1.1. *Wearable devices for fruit firmness estimation*

Advancements in wearable technology have revolutionized non-destructive testing methods, enhancing the efficiency and precision of fruit firmness estimation during harvesting and quality assessment. K. Peleg et al. pioneered a method utilizing vibration and acceleration transducers to evaluate the firmness of fruits and vegetables without causing damage (Peleg, 1997). Expanding on this foundation, Q. Lin et al. developed a wearable glove device capable of detecting and classifying agricultural products by measuring their curvature, color, and weight (Lin et al., 2018). Similarly, C. Pinto et al. introduced an intelligent glove equipped with sensors for pressure, color, and flexion, enabling real-time analysis of produce maturity and quality (Pinto et al., 2014).

Current wearable firmness estimation devices rely on low-resolution sensors, restricting their capability to capture detailed palpation and texture variations. Research on vision-based tactile sensing in wearable devices, which provides higher resolution and enhanced sensitivity, remains unexplored, despite its potential for more precise firmness assessment.

2.1.2. *Visuo-tactile-based devices for fruit firmness estimation*

Visuo-tactile devices have emerged as innovative tools for evaluating fruit firmness, a crucial factor in agricultural quality control. These devices facilitate non-destructive firmness measurements by seamlessly integrating advanced visual and tactile sensing technologies, significantly enhancing harvesting and storage practices. For instance, a visuo-tactile sensor designed to detect image variations during touch demonstrated remarkable efficacy, achieving an R^2 of 0.88 and an RMSE of 0.719 in assessing peach firmness (Ma et al., 2024). Similarly, a device employing a soft gripper inspired by the fin-ray effect combined tactile sensing with visual data processing, achieving R^2 values of 0.795 for tomatoes and 0.753 for nectarines (Lin et al., 2023). Another approach utilizes tactile predictive recognition for evaluating fruit hardness, delivering superior accuracy compared to conventional methods (Li et al., 2023).

Existing non-destructive fruit firmness estimation devices face several limitations. Mechanical devices like durometers depend on operator skill and are often fruit-specific, requiring additional assembly for different varieties (Harker et al., 1996). Acoustic and vibrational methods are affected by environmental factors such as temperature, humidity, and noise, compromising accuracy (Abbott et al., 1995). Optical devices, while advanced, are costly and inconsistent due to variations in fruit surface properties and environmental conditions like dust, moisture, and surface damage (Anjali et al., 2024). To overcome these challenges, we propose a novel wearable device that utilizes a vision-based tactile sensor for estimating fruit firmness. This device, designed for quality inspectors and farmers, is the first nondestructive approach to leverage off-the-shelf VBTS for mimicking human palpation. Worn on the thumb, it applies human-level pressure to the fruit surface, causing deformation in an elastomer. The RGB camera captures this deformation, and a deep learning model processes the data to predict firmness, enabling accurate and efficient ripeness grading.

3. Materials and methods

The proposed wearable device, illustrated in Fig. 1, was designed for the non-destructive, and real-time firmness assessment of 'Hayward' kiwifruit. This section describes the device's working principle and design, including integrating embedded systems for real-time processing. It also describes the plant material used, highlighting the creation of a dedicated dataset comprising tactile palpation recordings paired with penetrometer-based firmness measurements for model development. Furthermore, it also describes the proposed model and the implementation details that enabled accurate firmness estimation.

3.1. *Embodied human-device intelligence*

Embodied intelligence (Zhao et al., 2024), in the context of this paper, represents the seamless integration of human tactile abilities with advanced technological tools to achieve reliable and accurate firmness assessments. Haptic feedback, combining tactile and proprioceptive sensors, is fundamental to human touch. Tactile receptors detect changes in pressure, deformation, and contact area, while proprioceptive sensors in muscles and joints monitor finger movement and position. Together, these sensory inputs create a feedback loop that allows humans to adapt force precisely to material compliance without causing damage (Xu et al., 2020; Condon et al., 2014).

Despite this sophistication, human tactile assessments are inherently variable and often inconsistent. Repeated palpation, commonly used to verify firmness, can be destructive, particularly for delicate objects like fruits. Furthermore, humans may forget or misinterpret past tactile experiences, leading to unreliable evaluations. To overcome these challenges, our wearable device augments human capabilities, combining the adaptability of human touch with the precision and consistency of the device. The processing unit processes tactile information captured by the vision-based tactile sensor with reliability, ensuring accurate firmness assessments.

This synergy between human and device embodies intelligence by leveraging the operator's tactile instincts while standardizing and enhancing the assessment process. The device captures and processes

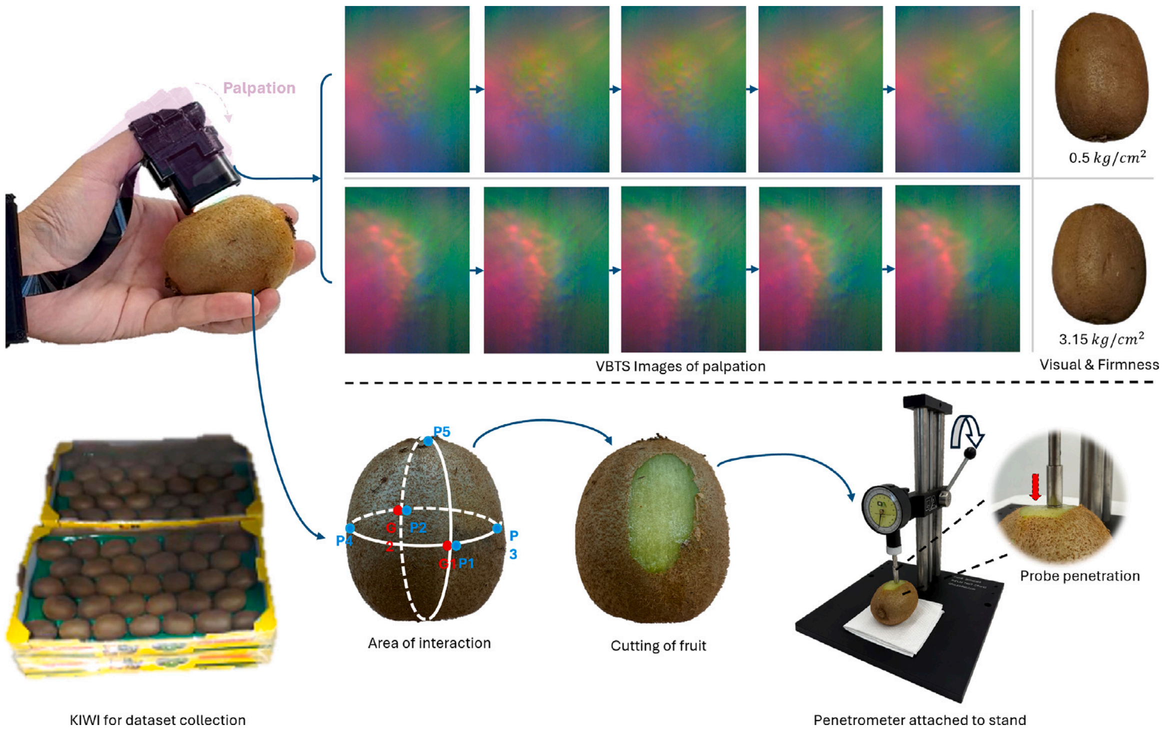

Fig. 3. Illustration of the dataset collection: The proposed device is used to record palpation data from each fruit. On the top right, two representative samples are shown—one soft (0.5 kg/cm^2) and one firm (3.15 kg/cm^2). Although both appear visually similar, their VBTS palpation signatures differ significantly, reflecting variations in firmness. At the bottom center, the blue points mark where each sample is palpated, and the red point indicates where the penetrometer reading is taken. Each Kiwi is then cut for penetrometer based ground truth measurements (Bottom right).

tactile information in real time, providing an objective assessment with a single palpation, unlike humans, who often rely on repeated comparisons to discern relative firmness.

3.2. Device design

The proposed device comprises two primary modules: a thumb module and a wrist module. The thumb module, worn on the user's thumb, incorporates a vision-based tactile sensor (DIGIT) at its tip (Lambeta et al., 2020), enabling high-resolution sensing of contact deformations during human palpation. The DIGIT sensor captures rich, real-time tactile data by observing the deformation of a soft elastomer surface through an internal camera, allowing for accurate recording of subtle textures, pressures, and natural interactions with the fruit. The wrist module contains the main control unit, a display, and a single user-friendly button for straightforward interaction. The device was designed symmetrically to accommodate different user preferences, allowing comfortable use on either hand. Fig. 2 illustrates the Computer-Aided Design (CAD) of our device.

To validate its design and functionality, a prototype was developed (Fig. 2). Both the wrist and thumb holders were 3D printed using black polylactic acid (PLA) filament. For additional comfort, the base of the wrist module was manufactured with a flexible NinjaFlex thermoplastic polyurethane (TPU). This was done to ensure a comfort fit on user's wrist. An adjustable soft strap was then attached to the base of wrist module and thumb module allowing the device to fit users of different sizes, making the thumb wearable both adaptable and secure during operation.

3.3. Embedded system

Local data processing is essential for our device to operate in remote locations where network connectivity is unreliable. To achieve this, the device was designed as a complete stand-alone unit for on-site inference. This approach ensures continuous operation and low latency.

At the core of the proposed system is the Radxa Zero 2 Pro single-board computer (SBC). This SBC is equipped with a quad-core ARM Cortex-A53 processor, a Mali-G31 MP2 GPU, up to 4 GB of LPDDR4 RAM, and eMMC storage, providing edge computational power for real-time tactile data processing. For user interaction, an OLED display was integrated into the system to provide visual feedback, and a button was included to initiate operations. The DIGIT VBTS (Lambeta et al., 2020), attached on the thumb module, captures high-resolution tactile palpation information, which are then processed locally on the SBC. Power is supplied through an external power source. All peripherals, including the display, buttons, VBTS, and power bank, are connected to the SBC via its General-Purpose Input/Output (GPIO) interface, establishing a centralized control system. Fig. 2 illustrates the hardware connections.

The operation workflow begins when the user presses the button, triggering the palpation process. The DIGIT sensor captures tactile images in the form of video, which are then processed by our proposed deep learning model deployed on the SBC. The model analyzes the palpation video in real time to predict the firmness of the object under examination. The firmness value is displayed on the OLED screen.

3.4. Plants materials

This section describes how the dataset was gathered, including selection of fruit, acquisition protocols, and labeling methods, all crucial for training and validating the model.

The Hayward Kiwifruit variety was selected for this study due to its widespread cultivation and its reputation for superior quality and appealing flavor (Garcia et al., 2012). This variety is known for its gradual decline in firmness during ripening, driven by physiological changes. Initially firm at harvest, the fruit softens over time as

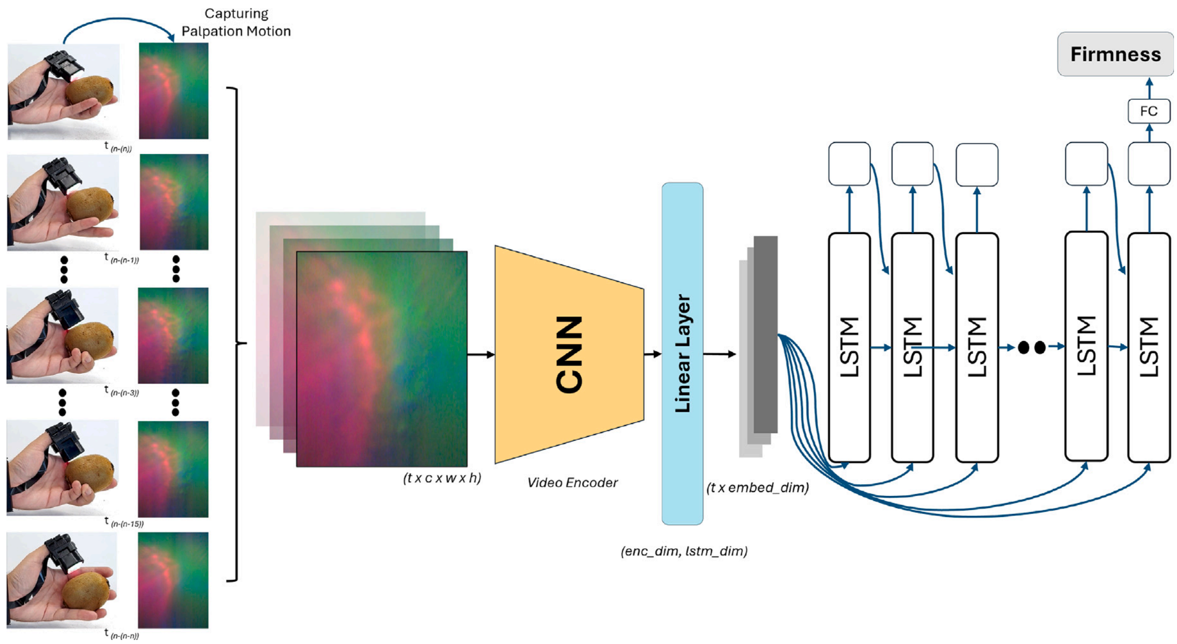

Fig. 4. Proposed architecture: VBTS palpation recordings, captured over n frames, are processed by a CNN-based video encoder to extract spatial features. A linear layer transforms these features at each time step into 1D spatial representations, subsequently fed into an LSTM network to model temporal dependencies. The final output is passed through a fully connected (FC) layer, generating the firmness prediction of the Kiwi.

cell adhesion weakens during cold storage. This softening accelerates in the later stages of ripening due to increased cell separation and greater plasticity of cell walls. These changes make accurate firmness assessment essential for maintaining fruit quality during post-harvest handling and storage (Harker and Hallett, 1994).

A total of 106 fruit samples were randomly selected for this study. Each fruit was palpated using the device at five distinct points (P1–P5) to capture localized variations in firmness. Ground truth firmness was measured using a QA Supplies penetrometer (Supplies, 2025). According to the standard protocol, approximately 1 mm of the fruit peel was removed to expose a flat flesh surface, and the probe was inserted perpendicularly into the fruit flesh to a depth of 7.9 mm (5/16 inch) over the course of a few seconds (Magness and Taylor, 1925; Li et al., 2016). Readings were considered invalid if the probe was inserted beyond or fell short of this marked depth. The penetrometer measured firmness in units of $\mathrm{kg/cm^2}$.

Firmness values (G1 and G2) for each fruit were measured at only two equatorial points, either P1–P2 or P3–P4, and the average of the two values was used as the ground truth. The P5 point, located near the stem, was not used for penetration-based firmness measurement, as creating a flat surface in this curved region would require the removal of fruit flesh, thereby violating the standard protocol. However, we intentionally included P5 in our sampling protocol for recording palpation only to ensure the dataset reflects real-world variability and improves the generalizability of our device across different fruit regions. Fig. 3 illustrates the dataset collection process and how the VBTS signatures differ between soft and hard samples.

This procedure yielded 530 unique pairs of VBTS palpation recordings and corresponding ground truth measurements. For each palpation recording, frames were selected from the start of contact until the end of contact. The number of frames per recording ranged from 32 to 96, with a mean of 44.32 (median 43). Ground truth firmness values ranged from 0.5 to 3.3 $\mathrm{kg/cm^2}$, with a mean of 1.4945 $\mathrm{kg/cm^2}$ (median 1.45 $\mathrm{kg/cm^2}$). While most samples clustered around the mean, a few outliers extended the range, especially very hard fruits. Finally, the dataset was split in an 80:20 ratio for training and testing, ensuring that the model is evaluated across the complete firmness spectrum—including atypical values.

3.5. Network architecture and implementation details

This study proposes a CNN-LSTM architecture for predicting fruit firmness from video sequences (Fig. 4). The model input consists of 16-frame sequences uniformly sampled from each video, ensuring consistent temporal coverage. Spatial features are extracted from individual frames using a pre-trained MobileNet v2 (Howard et al., 2017), which processes the spatial information within each frame. These extracted features are then passed to a single-layer LSTM with 128 hidden units, enabling the model to capture temporal dependencies across the sampled frames. This approach allows the CNN-LSTM architecture to effectively analyze palpation videos and estimate fruit firmness with high precision.

To ensure uniform temporal representation of the video data, frames were sampled at equal intervals. The sampling process divided the total number of frames, N, into consistent intervals corresponding to the desired number of sampled frames, n_{sample}. The step size is determined using the formula:

$$\text{step} = \max\left(1, \frac{N-1}{n_{\text{sample}}-1}\right)$$

The Huberloss (Huber, 1992) loss function optimizes the model's performance during training. It is defined as:

$$L(y,\hat{y}) = \begin{cases} 0.5\cdot(y-\hat{y})^2 & \text{if } |y-\hat{y}|\le\delta, \\ \delta\cdot|y-\hat{y}| - 0.5\cdot\delta^2 & \text{otherwise.} \end{cases}$$

Here, y represents ground truth, $\hat{y}$ represents predicted firmness, and δ represents the threshold at which the loss transitions from quadratic to linear, enhancing its robustness to outliers.

The training process employed the RMSprop optimizer and a CosineAnnealingLR scheduler to optimize the network's performance. Early stopping was applied after five epochs of no improvement to prevent overfitting. Pre-trained weights were fine-tuned over 1000 epochs with a learning rate and weight decay of 0.00005. Data augmentation techniques were employed to improve generalization, including random horizontal flipping, color jittering, and normalization.

The proposed model was implemented with Pytorch using the HuggingFace library on a machine with NVIDIA RTX 4090 GPU, CUDA Toolkit v11.0.221, and cuDNN v7.5. The effectiveness of the proposed

Table 1
Experimenting with the proposed model's various components and parameters as discussed in Section 4. In Exp# 1, the *Sample Duration* was experimented with for better temporal context. Exp# 2 explores the impact of *Loss Functions* for better accuracy, and lastly, in Exp# 3, the effect of optimizer was explored for convergence.

Sr	Model	Sample duration	Loss function	Optimizer	R^2 ↑
1	Baseline	8	SmoothL1	AdamW	0.70
2	Exp# 1	4	SmoothL1	AdamW	0.43
3		16	SmoothL1	AdamW	0.79
4	Exp# 2	16	MSEloss	AdamW	0.76
5		16	Huberloss	AdamW	0.85
6	Exp# 3	16	Huberloss	SGD	–
7		16	Huberloss	RMSprop	**0.89**

model in estimating firmness is comprehensively assessed and benchmarked against state-of-the-art (SOTA) methods using evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2).

4. Experimentation on CNN-LSTM model

The ablation study explores the impact of various configurations on the model's performance. The baseline configuration, utilizing a sample duration of 8 with the SmoothL1 loss function and AdamW optimizer, achieved an R^2 score of 0.70, serving as a reference point for further variations. Table 1 summarizes the experimentation.

Effect of Sample duration: Experiments were conducted to analyze the impact of varying sample duration on model performance while keeping the loss function (SmoothL1), optimizer (AdamW), and other hyperparameters constant. Results proved that decreasing the sample duration from 8 (baseline) to 4 reduced the R^2 value from 0.70 to 0.43. However, increasing the sample duration to 16 improved the R^2 value to 0.79. This indicates that a longer sample duration provides a better temporal context, improving model performance.

Impact of loss functions: The CNN-LSTM architecture was ablated incrementally. This section analyzes the model performance by changing loss functions. The baseline SmoothL1 loss was replaced with MSELoss and HuberLoss, while other training parameters were kept consistent, isolating the impact of the loss function. The performance decreased by 3.8%, with MSELoss resulting in a R^2 score of 0.76. For HuberLoss, the performance increased with a R^2 score of 0.85, showing a significant improvement of 7.59%. This suggests that HuberLoss effectively handles deviation compared to SmoothL1 and MSELoss. While MSELoss heavily penalizes large errors due to its quadratic nature, HuberLoss applys a linear penalty to large residuals, thus reducing the influence of outliers more effectively. Unlike SmoothL1, which has a fixed transition.

Effect of optimizers for better convergence: The impact of using different optimizers was explored while keeping the sample duration fixed at 16 and using HuberLoss. Although the model did not converge with the SGD optimizer, using RMSprop resulted in the highest R^2 score of 0.89. RMSProp adapts the learning rate for each parameter dynamically based on recent gradients. It maintains a moving average of the squared gradients, which helps stabilize updates and prevents large oscillations in the optimization process. This is particularly beneficial for datasets that contain outliers of varying scales, allowing them to converge more effectively.

Overall, the ablated model achieved a R^2 increased of 0.19 increment over the baseline. This improvement is attributed to the combination of high temporal information, Huberloss loss function for handling outliers, and RMSProp to maintain stability during training. It enhanced the model's capability to predict firmness from Palpation (video).

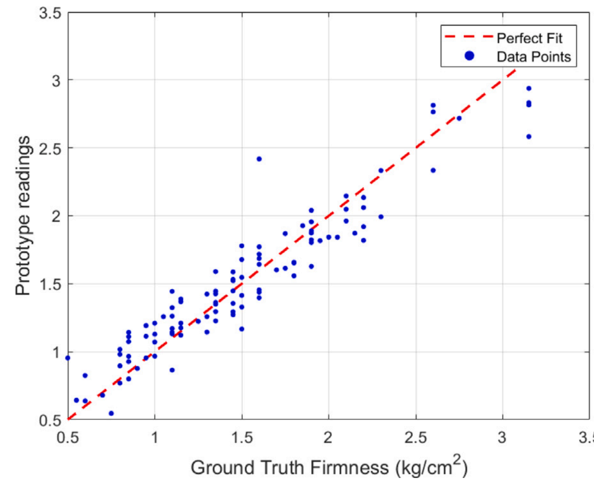

Fig. 5. Correlation between the prototype readings and reference firmness values with a Coefficient of Determination (R^2) score of 0.89.

5. Results and discussion

This section discusses the quantitative and qualitative results. Lastly, the comparison of our proposed model with SOTA architecture and the comparison of our proposed device all available commercials is also discussed.

5.1. Regression analysis

In the evaluation of our video regression model, we conducted a comprehensive residual analysis to assess the discrepancies between the model predictions and the observed ground truth values. The scatter plot, Fig. 6, of residuals versus ground truth demonstrates a generally random distribution of residuals across different ground truth values, suggesting that our model does not exhibit systematic bias across the range of outputs. This is further corroborated by the histogram of residuals, which shows a distribution centered around zero with most residuals tightly clustered within a narrow range, albeit with slight skewness. This distribution suggests that, while our model is generally accurate and unbiased, there remains some variability in prediction accuracy.

Fig. 5 illustrates the correlation between the firmness readings of the proposed model and the ground truth. The red dashed line represents the ideal 1:1 correlation, and the blue data points indicate the predictions, which closely follow the trend of the perfect fit line, demonstrating a strong positive correlation. A R^2 score of 0.89 was achieved.

5.2. Palpation explainability and interpretability analysis

To assess our model's ability to capture human palpation motions from DIGIT video sequences, we utilized Grad-CAM (Selvaraju et al., 2017) using the TorchCAM library (Fernandez, 2020). Fig. 7 showcases samples with varying firmness levels and corresponding spatial attention patterns. In each frame, the left image represents the input from the DIGIT sensor, while the right picture shows the overlay of the model's attention using Grad-CAM. This visualization technique highlights image regions most influential to the model's predictions. This allowed us to confirm whether the model aligns its focus with the DIGIT sensor motion during palpation.

For the first sample, the ground truth firmness was 2.75, and the model predicted 2.71, resulting in a residual of 0.03. In the second sample, the residual was only 0.02. Both samples exhibit strong spatial alignment between the model's attention and the palpation regions,

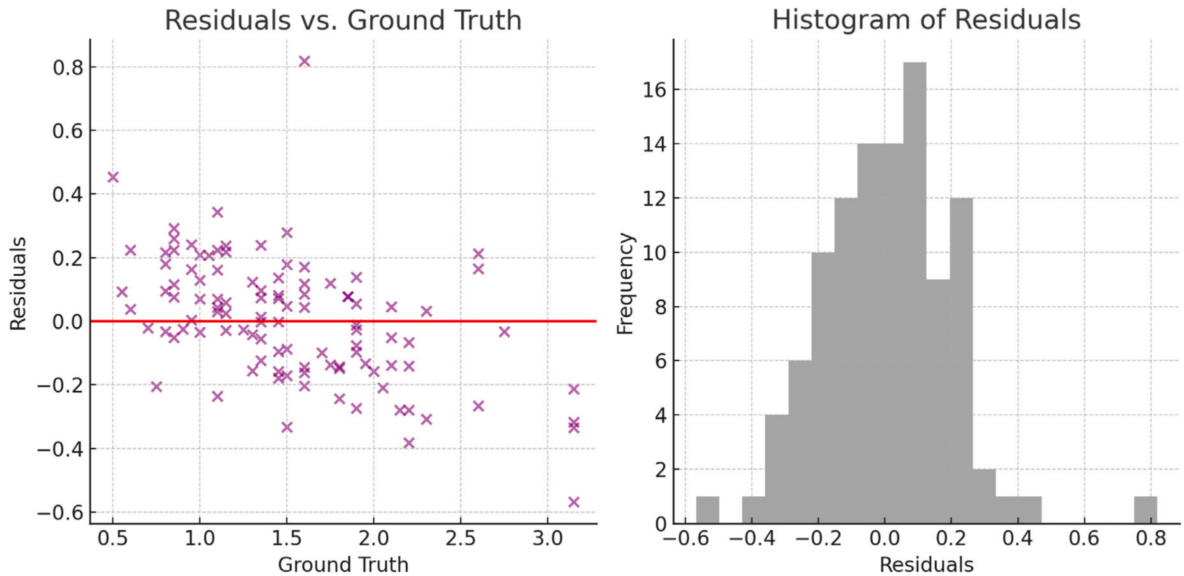

Fig. 6. Residual Analysis of our model **(Left)** Scatter plot of residuals demonstrates random distribution suggesting unbiased model estimation.**(Right)** The Histogram of Residuals shows residuals centered around zero, indicating general estimation accuracy with slight variability.

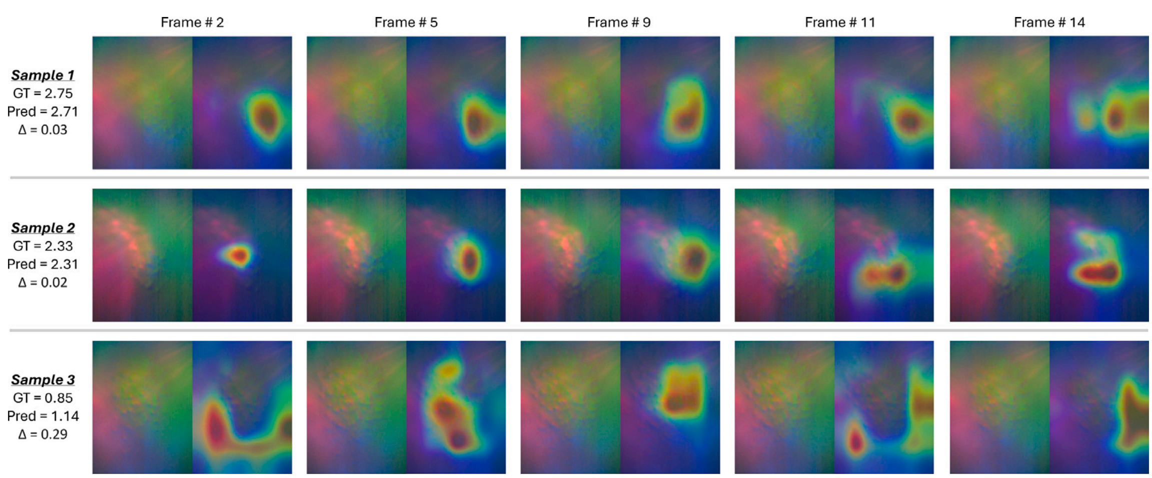

Fig. 7. Grad-CAM Visualizations of our proposed model: Each row represents a unique sample with corresponding ground-truth (GT) and predicted (Pred) firmness values, along with the residuals (Δ). The left image in each frame shows the input VBTS image, while the right image displays the model's attention overlay using Grad-CAM. The Grad-CAM visualizations highlight the spatial alignment between the model's attention and the palpation region on VBTS images. In Samples 1 and 2, strong alignment correlates with accurate estimation and low residuals. Conversely, Sample 3 shows poor alignment, resulting in higher residual error. The frames displayed correspond to the sequence's 2nd, 5th, 9th, 11th, and 14th positions.

indicating that the model effectively captures relevant palpation deformation for accurate firmness prediction. Conversely, in the third sample, the residual was 0.29. As shown in Fig. 7, the model's attention is less aligned with the palpation region in this sample, contributing to a higher residual error. Overall, the consistent attention patterns in well-aligned cases highlight the model's generalization capability and reliability in analyzing palpation dynamics, while misalignments suggest areas for further improvement in handling diverse firmness conditions.

Lastly, we analyze the impact of duration of palpation on prediction of firmness. Fig. 8 displays the relationship between the number of frames in a video and the prediction residuals of our regression model. The red line, which represents a smoothed trend of the residuals, exhibits a slight peak around 50 frames, indicating minor discrepancies in predictions in this range. However, as the number of frames increases beyond this point, the residuals trend towards zero, suggesting that the model predictions become more accurate or consistent with the actual values. This pattern implies that the model may handle longer sequences slightly better, possibly due to more comprehensive data within these videos. Nevertheless, the overall trend is relatively flat, confirming that the number of frames does not significantly affect the prediction residuals, indicating minimal influence on model performance across different video lengths. It is important to note that while the FPS of the DIGIT sensor was set to 30, we cannot confirm if this was maintained in real time since the actual duration of palpation was not recorded. As explained in Section 3.4, our dataset only contains frames of palpation, and the number of frames serves as an indicator of the duration of palpation.

This analysis confirms that the model focuses on relevant spatial and temporal features related to human palpation motions, supporting its ability to predict fruit firmness accurately.

5.3. Comparative analysis

The performance of the proposed CNN-LSTM model was evaluated through a comparative analysis with transformer based architectures, specifically TimesFormer (Bertasius et al., 2021) and VideoMAE (Tong et al., 2022), which represent the state-of-the-art in video-based tasks.

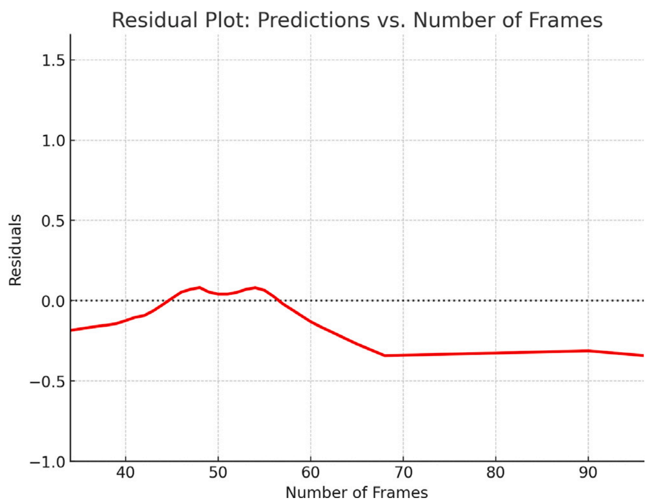

Fig. 8. Residual plot of proposed model firmness estimation vs. Number of Frames — Shows minimal discrepancies at around 50 frames, stabilizing in longer videos. Indicates minimal impact of frame count on model accuracy.

Table 2
Evaluation of the proposed model versus SOTA architectures on MSE, RMSE, R^2, and MAE metrics.

Sr	Model	MAE ↓	RMSE ↓	R^2 ↑
1	TimesFormer (Bertasius et al., 2021)	0.17	0.21	0.87
2	VideoMAE (Tong et al., 2022)	0.22	0.28	0.78
3	Ours *Baseline*	0.24	0.33	0.70
4	Ours *Improved*	**0.15**	**0.20**	**0.89**

To ensure a rigorous and equitable comparison, these transformer models were fine-tuned on the dataset used in this study, utilizing their pre-trained baseline weights to leverage the benefits of extensive prior training on large-scale datasets.

Our proposed model outperformed both transformers in terms of accuracy and error metrics. Specifically, our improved model achieved the lowest MAE of 0.15, RMSE of 0.20, and the highest R^2 value of 0.89. In contrast, TimesFormer achieved an MAE of 0.17 and RMSE of 0.21, with an R^2 of 0.87, while VideoMAE reported an MAE of 0.22, RMSE of 0.28, and R^2 of 0.78. These results indicate that our model provides a more precise and reliable prediction. Table 2 compares our proposed model with SOTA models.

5.4. Comparison with commercial devices

The proposed thumb-mounted device for the estimation of fruit firmness is compared to existing portable commercial devices in Table 3. These devices, widely used in the agricultural industry, serve as benchmarks for practical applications (see Fig. 9). Key attributes considered in the comparison include non-destructiveness, compactness, wearability, on-tree measurement capability, AI integration, wireless functionality, ease of operation, and cost-effectiveness. Non-destructiveness preserves fruit quality, compactness, and wearability, enhancing usability, and on-tree measurement capability allows immediate field data collection. AI integration improves accuracy, wireless functionality simplifies data management, ease of operation ensures accessibility for diverse users, and cost-effectiveness broadens adoption.

Table 3 details the evaluation of these attributes and reveals notable limitations in existing devices. Mechanical penetrometers (e.g., FT-series, GY series, FHT series) are compact, wireless, and low-cost but destructive, non-wearable, and lack on-tree measurement capability. Mechanical durometers (e.g., HPEII-Fff, Fruit Firm Meter, Firmness Tester) are compact, wireless, and support on-tree measurements but are not wearable and require additional apparatus for accuracy. Acoustic/vibrational devices vary: the Aweta AFS is non-destructive but bulky, non-wearable, and requires additional shape and weight measurements, while the MR-series is compact, wearable, and wireless but has accuracy limitations and is high-cost. Optical devices (e.g., F-750, NIR Case, DA-meter) are non-destructive, with the DA-meter being compact and wireless, while others are larger and may require assembly for different fruits.

In contrast, our device stands out by incorporating all key attributes—it is non-destructive, compact, wearable, supports on-tree measurements, and is wireless. It has no noted accuracy or operational limitations and employs a palpation-based method, making it unique among the compared devices. In addition, the cost is low.

Lastly, the global fruit firmness tester market was valued at 72.3 million USD in 2023 and is projected to reach 110.2 million USD by 2032 (More, 2024). This growth underscores the increasing demand for non-destructive firmness testers, highlighting the market potential and relevance of the proposed device as an innovative and practical solution for the agriculture industry.

6. Real world demonstration

To evaluate the device's practical utility in agriculture, we conducted on-tree firmness assessments of kiwifruit without detaching or damaging them—highlighting the non-destructive nature of our approach and its relevance for precision harvest decisions. As shown in Fig. 10, three kiwifruits were suspended from a tree for testing. The operator donned the wrist and thumb modules, completing setup in 30 ± 11 s. Firmness estimation involved gently grasping a fruit, initiating the process via a button press, and receiving real-time prompts and haptic feedback. The SBC verified sensor status, captured tactile data, ran the deep learning model, and displayed the predicted firmness within 18.25 ± 0.16 s. Including a 5-second display period, the full interaction loop lasted 23.98 ± 0.31 s. Model inference time was 3.82 ± 0.27 s, with a computational complexity of 5.01 GFLOPs for 16 frames (0.3128 GFLOPs/frame), measured using the FVCORE library (Research, Facebook AI, 2019).

The operator applied minimal force (2–10 N) with brief short palpation durations to minimize bruising. No wrist or thumb strain was reported, and the fruits showed no visible damage, validating the device's comfort and safety for on-tree use.

7. Limitations and future work

While the results are promising, this study is currently limited to only "Hayward" kiwifruit. Future work will aim to broaden the dataset by including a wider variety of fruits and vegetables in terms of size, shape, and weight, such as cucumbers, tomatoes, blueberries, and strawberries, and will involve multiple users to enhance device generalizability. Although the device is designed and demonstrated for non-destructive on-tree firmness estimation (see Section 6), the dataset collected and the proposed model were validated on harvested samples in the current study. Future work will focus on collecting a dataset of on-tree samples to assess the system's effectiveness in detecting subtle firmness changes during on-tree ripening.

At this stage, firmness estimation is based solely on palpation data. Future iterations will explore multimodal sensing by incorporating additional cues, such as applied force or smell, to improve ripeness prediction. Further enhancements in portability and efficiency are also planned through the integration of custom tactile sensors and dedicated PCB miniaturization.

Fig. 9. Comparison of the proposed wearable device and traditional fruit firmness testers. The proposed device integrates human dexterity with AI-driven tactile sensing, enabling non-destructive, real-time firmness estimation. In contrast, conventional methods—including mechanical (rupture and durometer), optical, and vibrational techniques are often invasive or influenced by environmental factors, limiting their adaptability and accuracy.

Table 3
Qualitative comparison of portable commercial devices used for fruit firmness detection.

Sr	Mechanism	Name	Non-Destructive	Compact	Wearable	On-tree	AI-Based	Wireless	Accuracy	Operation	Cost	Weight (g)	Size (mm)
1	Mechanical (Rupture Force)	FT-series (Supplies, 2024)	✗	✓	✗	✗	✗	✓	α	ω	$	500	112*59*24
2		GY series (Handpi, 2024)	✗	✓	✗	✗	✗	✓	α	ω	$	500	140*60*30
3		FHT series (Landtek, 2024)	✗	✓	✗	✗	✗	✓	α	ω	$	200	204*62*33
4	Mechanical (Deformation)	HPEII-Fff (Bareiss, 2024)	✓	✓	✗	✓	✗	✓	α	ω	$$$$$	300	190*70*40
5		Fruit Firm Meter (Turoni, 2024)	✓	✓	✗	✓	✗	✓	α	ω	$$	–	–
6		Firmness Tester (Aweta, 2024)	✓	✓	✗	✓	✗	✓	α	ω	$$	210–400	65*60*27 130*72*33
7	Acoustic/ Vibrational	Aweta AFS (Aweta, 2024)	✓	✗	✗	✗	✗	✗	ϕ	β	–	–	–
8		MR-series (MR-Series, 2024)	✓	✓	✓	✓	✗	✓	γ	–	$$$$	76	Wearable
10	Optical	F-750 (Felix, 2024)	✓	✗	✗	✓	✗	✓	ϕ	ϕ	$$$$$	1050	180*120*45
11		NIR Case (Sacmi, 2024)	✓	✗	✗	✗	✗	✗	ϕ	ω	–	NA	400*300*200
12		DA-meter (Store, 2024)	✓	✓	✗	✓	✗	✓	ϕ	ϕ	–	320	165*80*50
13	Vision-based tactile sensing	Ours	✓	✓	✓	✓	✓	✓	ϕ	ϕ	$	135	Wearable

Notes: In the 'Accuracy Limitations' column, symbols represent specific conditions: 'α' indicates that additional apparatus is necessary to enhance accuracy, 'γ' suggests that the accuracy of the measurement is affected by contact, and 'ϕ' signifies none of the aforementioned conditions apply. In the 'Operation Limitations' column, 'ω' denotes that additional assembly of components is required for handling different fruits, 'β' implies additional measurements related to the shape and weight of the fruit are necessary, and 'ϕ' indicates none of the mentioned limitations are applicable. In the 'Cost' column, prices are indicated as follows: $ = under $1,000; $$ = $1,000-$1,999; $$$ = $2,000-$2,999; $$$$ = $3,000-$4,999; $$$$$ = over $5,000.

8. Conclusion

Ripeness is a key determinant of fruit quality, influencing flavor, marketability, and waste reduction. However, assessing ripeness in fruits without clear visual cues is challenging, often necessitating destructive firmness measurements. The standard method, a penetrometer, involves plucking sample fruits, transporting them to a lab, cutting them open, and probing them—an inefficient, time-consuming process. If the sampled fruits are not representative, the results may mislead harvesting decisions and contribute to unnecessary waste. Humans naturally assess firmness through palpation, a non-destructive yet inconsistent approach that often requires multiple attempts, increasing the risk of bruising and subjective variability.

This work introduces a novel wearable system that merges human tactile interaction with vision-based tactile sensing and embedded deep learning. Operating on an edge computing platform, it bypasses cloud dependencies, ensuring reliable use in low-connectivity environments. By embedding a DIGIT sensor at the thumb tip, the device harnesses natural human dexterity, while the proposed model processes spatiotemporal features of palpation to provide rapid, accurate firmness predictions. The device comprises two modules: the thumb module, which houses a DIGIT sensor to capture real-time tactile images, and the wrist module, which contains the controller, display, and user interface. The operator holds the kiwifruit with four fingers and presses a button to initiate palpation. The embedded model then processes vision-based tactile sensing (VBTS) data to predict firmness, ensuring non-destructive evaluation in a real-world agricultural setting.

To validate the system, a Hayward kiwifruit dataset containing 530 unique palpation-firmness pairs was collected. The proposed model achieved an R^2 score of 0.89, demonstrating high accuracy. A real-world demonstration confirmed the feasibility of on-tree firmness estimation. No bruising was observed, and users reported the device as comfortable to wear.

The proposed device outperforms existing commercial fruit firmness testers by combining non-destructive, compact, wearable, and wireless features with AI integration. These findings pave the way for more efficient, consistent, and user-friendly quality assessments in the agricultural sector, addressing key limitations of current fruit firmness evaluation methods. This matters at scale: WWF estimates that about 15% of food is lost before it leaves the farm, largely due to poor harvest

Fig. 10. On-tree demonstration of the wearable device for real-time kiwifruit firmness assessment. This figure shows the device in a field setting, estimating firmness directly on the tree using palpation-based sensing. The battery-powered, wrist-worn design supports portable, non-destructive use. The setup demonstrates on-tree feasibility.

timing and limited testing tools (World Wide Fund for Nature, 2021). By providing quick, accurate firmness readings in the orchard, the device can curb premature picking, reduce bruising, and cut pre-harvest waste, supporting more sustainable production.

Furthermore, cost remains the chief barrier to adoption, so recent reviews advocate interim, low-cost technologies to keep expenses manageable (Oliveira et al., 2021; Duckett et al., 2018). In line with this need, the fruit firmness tester market was valued at USD 72.3 million in 2023 and is projected to reach USD 110.2 million by 2032 (More, 2024), indicating growing demand for affordable, non-destructive solutions and underscoring the market potential of our device.

CRediT authorship contribution statement

Mashood M. Mohsan: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Basma B. Hasanen:** Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Taimur Hassan:** Writing – review & editing. **Lakmal Seneviratne:** Visualization, Supervision, Resources, Project administration, Funding acquisition. **Irfan Hussain:** Writing – review & editing, Visualization, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This publication is based upon work supported by the Khalifa University of Science and Technology, United Arab Emirates under Award No. RC1-2018-KUCARS.

Data availability

The dataset will be released after acceptance.

References

Abbott, Judith A., 1999. Quality measurement of fruits and vegetables. Postharvest Biol. Technol. 15 (3), 207–225, Publisher: Elsevier.
Abbott, J.A., et al., 1995. Nondestructive sonic firmness measurement of apples. Trans. ASAE 38 (5), 1461–1466, Publisher: American Society of Agricultural and Biological Engineers.
Ahmadi, Ebrahim, 2018. Bruise susceptibilities of kiwifruit as affected by impact and fruit properties.. Res. Agric. Eng. http://dx.doi.org/10.17221/57/2011-RAE.
Anjali, et al., 2024. State-of-the-art non-destructive approaches for maturity index determination in fruits and vegetables: Principles, applications, and future directions. Food Prod. Process. Nutr. 6 (1), 56, Publisher: Springer.
Aweta, 2024. Aweta - sorting and packing solutions for fruit and vegetables. URL https://www.aweta.com/en/.
Bareiss, 2024. Fruit firmness tester - HPE III fff - bareiss north america. URL https://www.bareiss-testing.com/product/fruit-firmness-tester-hpe-iii-fff/.
Barrett, Diane M., Garcia, Elisabeth, and, Jo Ellen Wayne, 1998. Textural modification of processing tomatoes. Crit. Rev. Food Sci. Nutr. 38 (3), 173–258. http://dx.doi.org/10.1080/10408699891274192, Publisher: Taylor & Francis _eprint:.
Bertasius, Gedas, Wang, Heng, Torresani, Lorenzo, 2021. Is space-time attention all you need for video understanding? In: ICML, vol. 2, p. 4, Issue: 3.
Cirilli, Marco, et al., 2016. On-field monitoring of fruit ripening evolution and quality parameters in olive mutants using a portable NIR-AOTF device. Food Chem. 199, 96–104, Publisher: Elsevier.
Condon, Melia, et al., 2014. Differential sensitivity to surface compliance by tactile afferents in the human finger pad. J. Neurophysiol. 111 (6), 1308–1317, Publisher: American Physiological Society Bethesda, MD.
Duckett, Tom, et al., 2018. Agricultural robotics: the future of robotic agriculture. arXiv preprint arXiv:1806.06762.
Felix, 2024. F-750 produce quality meter - felix instruments. URL https://felixinstruments.com/food-science-instruments/nir-spectroscopy/f-750-produce-quality-meter/.
Fernandez, François-Guillaume, 2020. TorchCAM: class activation explorer. URL https://github.com/frgfm/torch-cam.
Garcia, Coralia V., et al., 2012. Kiwifruit flavour: A review. Trends Food Sci. Technol. 24 (2), 82–91. http://dx.doi.org/10.1016/j.tifs.2011.08.012, URL https://www.sciencedirect.com/science/article/pii/S0924224411002445.
Handpi, 2024. Handpi - the force gauge and torque gauge specialists - force & torque testing solutions - worldwide. URL http://www.handpi.com/en/Productshow.asp?ArticleID=303.
Harker, F.R., Hallett, Ian C., 1994. Physiological and mechanical properties of kiwifruit tissue associated with texture change during cool storage. J. Am. Soc. Hortic. Sci. http://dx.doi.org/10.21273/JASHS.119.5.987.
Harker, F.R., Maindonald, J.H., Jackson, P.J., 1996. Penetrometer measurement of apple and kiwifruit firmness: operator and instrument differences. J. Am. Soc. Hortic. Sci. 121 (5), 927–936, Publisher: American Society for Horticultural Science.
Howard, Andrew G., et al., 2017. MobileNets: Efficient convolutional neural networks for mobile vision applications. _eprint: 1704.04861, URL https://arxiv.org/abs/1704.04861.
Huang, Yuping, Lu, Renfu, Chen, Kunjie, 2018. Prediction of firmness parameters of tomatoes by portable visible and near-infrared spectroscopy. J. Food Eng. 222, 185–198, Publisher: Elsevier.
Huber, Peter J., 1992. Robust estimation of a location parameter. In: Breakthroughs in Statistics: Methodology and Distribution. Springer, pp. 492–518.

Jantra, Chaiya, et al., 2018. Development of a handheld precision penetrometer system for fruit firmness measurement. Postharvest Biol. Technol. 144, 1–8, Publisher: Elsevier.

Khalifa, Smail, Komarizadeh, Mohammad Hassan, Tousi, Behrouz, 2011a. Review article usage of fruit response to both force and forced vibration applied to assess fruit firmness- a review.

Khalifa, Smail, Komarizadeh, Mohmmad Hasan, Tousi, Behrouz, 2011b. Usage of fruit response to both force and forced vibration applied to assess fruit firmness-a review. Aust. J. Crop. Sci. 5 (5), 516–522, Publisher: Southern Cross Publishers Lismore, NSW.

Khan, Asim, et al., 2023. Tomato maturity recognition with convolutional transformers. Sci. Rep. 13 (1), 22885, Publisher: Nature Publishing Group UK London.

Lambeta, Mike, et al., 2020. DIGIT: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation. http://dx.doi.org/10.1109/LRA.2020.2977257.

Landtek, 2024. Fruit hardness tester FHT-05, FHT-15, FHT-1122 catalog. URL https://www.landtek.cn/FruitHardnessTester_FHT-05_FHT-15_FHT-1122_Catalog.pdf.

Li, Huajia, et al., 2016. Kiwifruit firmness: Measurement by penetrometer and non-destructive devices. Postharvest Biol. Technol. 120, 127–137, Publisher: Elsevier.

Li, Song, et al., 2023. Assessing fruit hardness in robot hands using electric gripper actuators with tactile sensors. http://dx.doi.org/10.1016/j.sna.2023.114843.

Lin, Chien-Der, et al., 2018. Agricultural product detection device, agricultural product management system and method thereof.

Lin, Jiahao, et al., 2023. Non-destructive fruit firmness evaluation using a soft gripper and vision-based tactile sensing. Comput. Electron. Agric. http://dx.doi.org/10.1016/j.compag.2023.108256.

Ma, Chan, Ying, Yibin, Xie, Lijuan, 2024. Visuo-tactile sensor development and its application for non-destructive measurement of peach firmness. Comput. Electron. Agric. http://dx.doi.org/10.1016/j.compag.2024.108709.

Magness, John Robert, Taylor, George F., 1925. An improved type of pressure tester for the determination of fruit maturity.

Mazhar, M., et al., 2016. Comparison of firmness meters for measuring 'hass' avocado fruit firmness. http://dx.doi.org/10.17660/ACTAHORTIC.2016.1119.22.

Mohsan, Mashood M., et al., 2025. SwishFormer for robust firmness and ripeness recognition of fruits using visual tactile imagery. Postharvest Biol. Technol. 225, 113487. http://dx.doi.org/10.1016/j.postharvbio.2025.113487, URL https://www.sciencedirect.com/science/article/pii/S0925521425000997.

More, Anuradha B., 2024. Fruit firmness tester market: Global industry analysis, size, share, growth, trends, and forecast. URL https://dataintelo.com/report/global-fruit-firmness-tester-market.

MR-Series, 2024. MRI series. URL https://www.ava.co.jp/yubi.

Nazir, Mian Faisal, et al., 2024. Kiwifruit in the omics age: Advances in genomics, breeding, and beyond. Plants http://dx.doi.org/10.3390/plants13152156.

Oliveira, Luiz F.P., Moreira, António P., Silva, Manuel F., 2021. Advances in agriculture robotics: A state-of-the-art review and challenges ahead. Robotics 10 (2), 52, Publisher: MDPI.

Peleg, Kalman, 1997. Method and equipment for measuring firmness of fruits and vegetables.

Pinto, Ciro Martins, et al., 2014. BR102014013727a2 - wearable glove for analysis and classification of agricultural products - google patents. URL https://patents.google.com/patent/BR102014013727A2/en?oq=BR102014013727A2.

Rehman, Mobeen Ur, Eesaar, Hassan, Abbas, Zeeshan, Seneviratne, Lakmal, Hussain, Irfan, Chong, Kil To, 2024. Advanced drone-based weed detection using feature-enriched deep learning approach. Knowledge-Based Systems 305, 112655.

Research, Facebook AI, 2019. Fvcore: A collection of common code for deep learning research. URL https://github.com/facebookresearch/fvcore.

Rivera, Sebastian, et al., 2023. Instrumental mechanical parameters related to hand-feel touch firmness of blueberries. Postharvest Biol. Technol. 205, 112530. http://dx.doi.org/10.1016/j.postharvbio.2023.112530, URL https://www.sciencedirect.com/science/article/pii/S0925521423002910.

Sacmi, 2024. NIR inspection systems by sacmi. URL https://sacmi.com/en-US/corporate/news/5819/NIR-inspection-systems-by-Sacmi.

Selvaraju, Ramprasaath R., et al., 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 618–626.

Store, T.T., 2024. DA meter. URL https://www.trturoni.com/en/fruit-veg-ripeness-quality-control/da-meter/da-meter.5.html.

Supplies, Q.A., 2024, FT 011 Penetrometer, URL https://qasupplies.com/ft-011-penetrometer/.

Supplies, Q.A., 2025. FT 327 penetrometer. URL https://qasupplies.com/ft327-penetrometer/. (Accessed: 09 April 2025).

Tong, Zhan, et al., 2022. Videomae: Masked autoencoders are data-efficient learners for self-supervised video pre-training. Adv. Neural Inf. Process. Syst. 35, 10078–10093.

Turoni, 2024. TR Turoni 53225 fruit firm hardness tester for Kiwi, Blueberries, Cherries and more. URL https://tinyurl.com/tn96dym4.

Valiente, L.D., Parco, K.M.R., Sangalang, G.C.P., 2021-10. Non-destructive image processing analysis for defect identification and maturity detection on avocado fruit. In: 2021 5th International Conference on Communication and Information Systems. ICCIS, pp. 175–179. http://dx.doi.org/10.1109/ICCIS53528.2021.9645970.

Voss, Monica, et al., 2024. Unlocking the potential of agrifood waste for sustainable innovation in agriculture. Recycling http://dx.doi.org/10.3390/recycling9020025.

World Wide Fund for Nature, 2021. Over 15% of food is lost before leaving the farm - WWF report. URL https://www.wwf.eu/?4049841/fifteen-per-cent-of-food-is-lost-before-leaving-the-farm-WWF-report.

Xu, Chang, et al., 2020. Tactile exploration strategies with natural compliant objects elicit virtual stiffness cues. IEEE Trans. Haptics 13 (1), 4–10. http://dx.doi.org/10.1109/TOH.2019.2959767.

Yuan, Wenzhen, Srinivasan, Mandayam A., Adelson, Edward H., 2016. Estimating object hardness with a GelSight touch sensor. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 208–215. http://dx.doi.org/10.1109/IROS.2016.7759057.

Yuan, Wenzhen, Zhu, Chenzhuo, et al., 2017. Shape-independent hardness estimation using deep learning and a GelSight tactile sensor. In: 2017 IEEE International Conference on Robotics and Automation (ICRA). pp. 951–958. http://dx.doi.org/10.1109/ICRA.2017.7989116.

Zhao, Zikai, et al., 2024. Exploring embodied intelligence in soft robotics: A review. Biomimetics 9, http://dx.doi.org/10.3390/biomimetics9040248, URL https://www.mdpi.com/2313-7673/9/4/248/pdf?version=1713539955.