COMPUTER VISION-BASED GRASP DETECTION FOR A METAMATERIAL SOFT GRIPPER IN ROBOTIC VEGETABLES HARVESTING
/
DETECTAREA PRINDERII UNUI GRIPPER SOFT DIN METAMATERIALE UTILIZAT LA RECOLTAREA ROBOTIZATĂ A LEGUMELOR, FOLOSIND VEDERE ARTIFICIALĂ

Florin Bogdan MARIN [1)], Mihai Gabriel MATACHE*[2)], Mihaela MARIN*[1)], Gheorge GURAU[1)], Robert CRISTEA[2)], Andrei TĂNASE[3)]
[1)] "Dunarea de Jos" University of Galati / Romania;
[2)] INMA Bucharest / Romania;
[3)] INCDBH Ștefănești / Romania
Tel: +40 727 957 693; E-mail: matache@inma.ro
Tel: +40 741 681 400; E-mail: mihaela.marin@ugal.ro
DOI: https://doi.org/10.35633/inmateh-77-71

Keywords: *CNN, metamaterial, soft grippers, smart horticulture*

ABSTRACT
This paper presents a computer vision-based methodology for evaluating the grasping performance of a soft robotic gripper fabricated from mechanical metamaterials, designed specifically for fruit and vegetables harvesting applications. Due to the fragile nature of fruits such as tomatoes or strawberries, the ability to assess and control the deformation of the gripper during interaction is critical to avoid damage while ensuring a secure grasp. A deep learning approach is proposed, leveraging convolutional neural networks (CNNs) to classify grasp outcomes from visual input. The model is trained on a custom dataset of images captured during robotic harvesting trials and optimized to detect subtle variations in gripper shape and fruit contact. The integration of soft metamaterial-based grippers with computer vision algorithms enables a robust, non-invasive grasp assessment pipeline, contributing toward fully autonomous and adaptive fruit-picking robots. The proposed method achieved an accuracy of 94.0% for correct grasps, 91.5% for failed grasps, and 95.9% for no-object cases, with an average inference time of 87 ms (ranging from 75 to 98 ms).

REZUMAT
Această lucrare prezintă o metodologie bazată pe viziune artificială pentru evaluarea performanței de prindere a unui griper robotic moale, fabricat din metamateriale mecanice, conceput special pentru aplicații de recoltare a fructelor și legumelor. Având în vedere natura fragilă a fructelor, cum ar fi roșiile, capacitatea de a evalua și controla deformarea griperului în timpul interacțiunii este esențială pentru a evita deteriorarea acestora și pentru a asigura o prindere sigură. Propunem o abordare bazată pe învățare profundă, care utilizează rețele neuronale convoluționale (CNN) pentru a clasifica rezultatul prinderii pe baza informațiilor vizuale. Modelul este antrenat pe un set de date personalizat, constând în imagini capturate în timpul testelor de recoltare robotică, fiind optimizat pentru a detecta variații subtile în forma griperului și în contactul acestuia cu fructul. Integrarea griperelor moi realizate din metamateriale cu algoritmi de viziune artificială permite dezvoltarea unui sistem robust și neinvaziv de evaluare a prinderii, contribuind la crearea unor roboți autonomi și adaptivi pentru culesul fructelor. Metoda propusă a atins o acuratețe de 94,0% pentru prinderile corecte, 91,5% pentru prinderile eșuate și 95,9% pentru cazurile fără obiect, cu un timp mediu de procesare de 87 ms (între 75 și 98ms).

INTRODUCTION
Fruit harvesting represents a critical, yet often labour-intensive and costly, component of the agricultural supply chain. Traditionally, this process has relied heavily on manual labour, rendering it vulnerable to fluctuations in labour markets, rising costs, and the seasonal availability of workers. Furthermore, the repetitive and physically demanding nature of harvesting tasks frequently leads to operator fatigue and inconsistent efficiency. These challenges have fuelled increasing demand for automation solutions in agriculture, particularly within the fruit harvesting sector *(Bharad et al, 2024*). Automation not only promises reduced operational costs and improved efficiency, but also addresses labour shortages and ensures consistent product quality.

Despite the significant benefits of harvesting automation, its implementation is particularly challenging for delicate fruits. Unlike the harvesting of grains or more robust crops, fruits such as strawberries, tomatoes, apples, and peaches are susceptible to mechanical damage—including bruising, tearing, or crushing—if not handled with extreme care. This fragility imposes strict requirements on robotic harvesting systems, particularly on the design and performance of end-effectors (grippers). Experimental studies on kiwifruit harvesting robots demonstrated that the optimization of gripping force, rotation angle, and rotation speed is crucial to achieve high success rates and low fruit damage, emphasizing the importance of precise parameter tuning in robotic harvesting (*He et al., 2023*). A robotic gripper must exert sufficient force to detach the fruit, yet remain gentle enough to avoid damaging it. Additionally, it must navigate complex, foliage-rich environments, often with irregular fruit positioning, while avoiding harm to plants or adjacent fruits. Natural variations in fruit size, shape, ripeness, and spatial distribution further increase the complexity of the task, necessitating adaptable and intelligent robotic systems. A recent review provides a comprehensive analysis of robotic harvesting technologies for vegetables of the *Solanaceae* family, highlighting various fruit detachment principles and specialized grippers tested in greenhouses and solariums, with particular focus on adapting to the morphological and structural characteristics of the target species (*Matache et al., 2025*).

Inspired by biological structures, soft robotic grippers offer an adaptive grip and uniform pressure distribution, making them ideal for handling irregularly shaped or fragile items (*Su et al, 2021*). A major advancement in the development of soft grippers is the integration of metamaterials. Metamaterials are engineered substances with unconventional mechanical properties, derived not from their chemical composition but from their microstructural geometries. By manipulating their internal architecture, metamaterials can be designed to exhibit unique mechanical behaviours, such as tuneable stiffness, responsiveness to external stimuli (e.g., light, heat, magnetic fields), or even auxetic behaviour (expanding laterally when stretched) (*Sui et al 2022*). Incorporating metamaterials into soft grippers enables the creation of highly adaptable and functional gripping systems capable of modulating their mechanical properties in real time to match the demands of specific harvesting tasks. This development opens new frontiers for safe and precise fruit manipulation. However, even the most advanced metamaterial-based soft gripper would be ineffective without accurate perception of its environment (*Visentin et al, 2023*). This is where computer vision becomes effective (*Tang et al, 2020*). Computer vision systems empower robotic platforms to perceive (*Liu et al, 2024*), identify, and localize fruits, as well as to determine the optimal position and orientation for successful gripping (*Elfferich et al, 2022*). Such systems facilitate the detection of ripe fruits, estimation of their 3D positions in space, identification of grasping points, and monitoring of the interaction between the gripper and the fruit (*Li et al, 2022*). Without robust computer vision, a harvesting robot would effectively operate blindly, leading to low success rates and increased fruit damage. Consequently, the intelligent integration of computer vision with metamaterial-based soft grippers is essential to the realization of truly autonomous and efficient fruit harvesting systems (*Ge et al, 2019)*.

Soft grippers represent a class of robotic end-effectors constructed from flexible and elastic materials such as silicone, rubber, or thermoplastic polymers. Soft grippers have gained attention in agricultural robotics due to their ability to operate in unstructured environments and handle delicate objects gently—qualities that are essential for fragile crop harvesting (*Navas et al., 2024*). The fundamental operating principle of soft grippers is based on the controlled deformation of their structure. Unlike traditional rigid grippers that rely on precise contact and controlled force, soft grippers use material compliance to conform to the shape of the manipulated object. This intrinsic property provides significant advantages, especially in applications involving fragile, irregular, or deformable objects such as fruits (*Nguyen et al, 2023*). However, their inherent compliance and deformation during grasping add complexity to the task of assessing whether a grasp was successful. This deformation can be actuated by various mechanisms, including pneumatic inflation or deflation of internal chambers, hydraulic systems, tendon-driven actuation, or smart materials that change shape in response to external stimuli (e.g., temperature or magnetic fields). Regardless of the actuation method, the goal is to allow the gripper to wrap around the object, distributing contact pressure evenly and minimizing stress points that could cause damage (*Tawk et al, 2022*). Materials used in soft gripper construction are selected for their flexibility and elasticity *(Huang et al, 2022*). Silicone, in particular, is widely adopted due to its chemical inertness, resistance to extreme temperatures, and moldability into complex shapes. Soft gripper designs range from single-fingered and multi-chamber tubular structures to more complex architectures inspired by biological organisms, such as tentacles or elephant trunks (*Hegde et al, 2024*). Many designs incorporate internal structuring or stiffness patterns to control deformation modes and optimize grasping behaviour (*Wang et al, 2023*).

A notable trend is the integration of embedded sensors into the soft material, creating sensorized grippers capable of measuring grip force, contact pressure, or even surface texture. These sensory capabilities provide essential feedback for adaptive control systems. In agriculture, soft grippers made off metamaterials are regarded as a promising solution for automating the harvesting of delicate fruits. Their ability to manipulate produce without causing mechanical damage is critical for maintaining quality and reducing post-harvest losses. Specific applications include harvesting strawberries, raspberries, tomatoes, apples, and other soft-skinned or irregularly shaped fruits. Studies have demonstrated the efficacy of soft grippers in minimizing bruising and other defects associated with mechanical handling. For example, soft grippers have been successfully tested in the context of apple harvesting, exhibiting reduced fruit damage and improved reliability in diverse field conditions. Metamaterials are artificial, engineered materials that derive their unique properties not from the chemical composition of their constituents, but from their microscopic or macroscopic geometric structure (*Pyo et al, 2024*). This precisely designed, repetitive architecture allows them to exhibit unusual physical properties that are not typically found in natural materials (*Mohammadi et al, 2023*). The concept of metamaterials was initially explored in the fields of optics and electromagnetism, but has since expanded into other disciplines, including mechanics, where mechanical metamaterials have emerged (*Shintake et al, 2018*).

The integration of metamaterials into soft grippers represents a promising research frontier. By embedding metamaterial structures into the gripper design, it becomes possible to obtain custom mechanical behaviour and enhanced functionality. For instance, a soft gripper (*Kaur et al, 2019*) can be designed with metamaterial segments that adapt their stiffness in real-time based on the required gripping force, or with auxetic structures that better conform to the fruit's surface. This integration supports the development of intelligent grippers capable of dynamically responding to harvesting environments, optimizing grasping strategies, and minimizing fruit damage. Flexible metamaterials can significantly enhance the performance of soft robots by enabling complex motion behaviours that are inherently programmed into monolithic structural systems (*Zhou et al, 2017*).

Grasp detection plays a central role in ensuring operational performance. A harvesting cycle is only successful if the fruit is both correctly identified and securely detached from the plant. Any false detection—such as an apparent grasp without actual retention of the fruit—leads to inefficient operation, increased cycle times, and potential damage to adjacent produce. Moreover, in the absence of real-time grasp validation, the robot cannot correct failed attempts, leading to suboptimal performance in unstructured environments. Detecting the quality of the grasp also helps prevent mechanical damage to fragile produce by avoiding excessive or asymmetric forces that could bruise or deform the fruit (*Seo et al., 2024*).

Given that soft grippers made from metamaterials exhibit programmable stiffness and nonlinear deformation characteristics, the ability to monitor their shape evolution during interaction is equally essential. Observing how the gripper deforms provides implicit information about the position, size, and firmness of the fruit being handled. For instance, symmetrical deformation may indicate a correct and centred grasp, while uneven shape changes can signal slippage or off-axis contact. Such information can be extracted using vision-based systems, eliminating the need for embedded force or pressure sensors, thus reducing system complexity, and improving robustness (*Grady et al., 2022*).

Furthermore, understanding the deformation behaviour of metamaterial structures under load contributes to more accurate modelling of the object–gripper interaction, enabling contact-aware adaptive control of the gripping strategy in real time (*Xie et al., 2024*). This is particularly relevant when dealing with fruits that vary significantly in ripeness, texture, and geometry. Adaptive control algorithms can exploit this feedback to adjust the gripping strategy in real time, improving both precision and safety. Consequently, combining grasp detection with deformation monitoring not only enhances the mechanical performance of the system but also facilitates the development of intelligent robotic platforms capable of operating autonomously in dynamic agricultural environments.

The aim of this study is to evaluate a metamaterial-based soft gripper for vegetable harvesting and to develop a computer vision method that uses a convolutional neural network (CNN) to determine whether a grasp is correct, failed, or occurs without an object, based on the observed deformation of the gripper.

MATERIALS AND METHODS

The robotic platform specifically designed for the autonomous harvesting of delicate fruits is fitted with a soft gripper. The system integrates a mobile robotic base, a multi-degree-of-freedom robotic arm, a soft gripper fabricated using mechanical metamaterials, and an embedded vision module. Each component has been carefully selected and configured to address the challenges associated with handling soft, deformable, and irregularly shaped agricultural products such as tomatoes fruits.

The key Components and Functional Elements which compose the robotic platform, presented in Figure 1 are the following:

(1) Integrated Vision System

Mounted adjacent to the gripper, the onboard camera serves as the primary sensory input for the computer vision module. It enables real-time fruit detection, grasp point selection, and grasp validation. The placement of the camera ensures minimal occlusion during manipulation tasks and provides a clear line of sight for object recognition and tracking. The vision system consisted of a Runex S700 Webcam RGB camera with a resolution of [800X600] pixels and a frame rate of 30 fps, positioned at a distance of 100 mm from the gripper.

(2) Soft Gripper Based on Mechanical Metamaterials

The end-effector is constructed using flexible materials with engineered microstructures, providing it with adaptive compliance and distributed pressure during grasping. Unlike conventional rigid grippers, this soft gripper can conform to the geometry of the target object, thereby minimizing localized stress and reducing the risk of mechanical damage to the fruit.

(3) Joint Actuation Modules

The robotic arm is actuated by a series of servo motors, allowing controlled articulation of each joint. This enables the gripper to manoeuvre in complex environments, navigate foliage, and align optimally with fruits situated at varying orientations and positions. The robotic arm used for manipulation had 4 degrees of freedom and a maximum payload capacity of 0.5 kg, with a working envelope of [900] × [300] × [500] mm.

The platform has the following degrees of Freedom and Motion Axes:

- Z-axis – Vertical actuation of the entire robotic structure, allowing adjustment of working height to reach fruits positioned at different levels within the canopy.
- V-axis – Vertical pitch rotation of the robotic arm relative to its base support, providing flexibility in approach angles.
- W-axis – Horizontal yaw rotation of the arm base, enabling the system to access targets across a wide field of view without needing to reposition the mobile platform.
- T-axis – Terminal rotation at the gripper level, facilitating precise orientation for a stable and safe grasp.

The robotic system was mounted on a wheeled base for further development of semi-autonomous mobility within a greenhouse. The modular architecture allows for easy upgrades, including the replacement of the gripper or vision system depending on crop type or harvesting task.

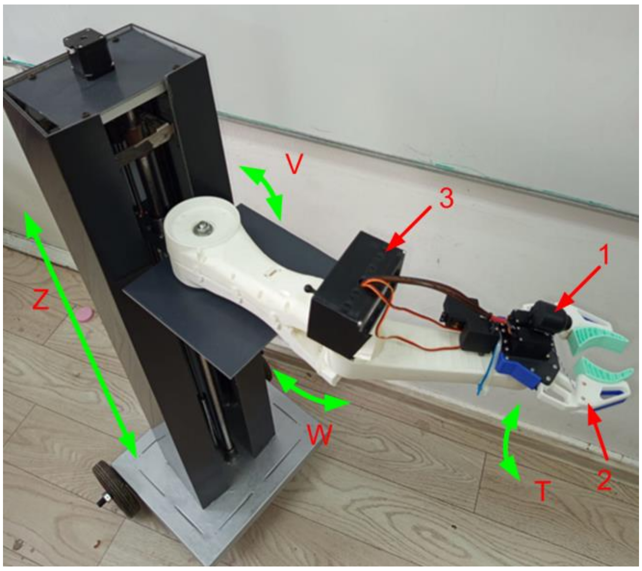

Fig. 1 - Experimental Robot Platform - Technical Equipment for fruit harvesting
1. Runex S700 web camera; 2. Gripper; 3. Gripper controller;
Z vertical movement V first rotation W second rotation T- third rotation

A soft insert for the gripper was designed (Figure 2), with carefully defined spacing between the ribs of each finger. The component was manufactured from TPU 85A, a flexible thermoplastic polyurethane known for its excellent compliance and resistance when handling fragile loads. The insert material exhibits a Shore A hardness of 85, a tensile strength of 34 MPa, and an elongation at break of 600 %, providing sufficient elasticity for compliant fruit handling. The insert, mounted on the inner face of the gripper in direct contact with the fruits, measures 23 mm in width at the base, 65 mm in length, and 25 mm in thickness. It contains 11 ribs, unevenly distributed between the two edges, with an average spacing of approximately 3.2 mm.

Each rib has a rounded cross-section, with a height of 2 mm above the finger surface, ensuring both flexibility and slip resistance. The contact surface of the ribs is slightly textured to increase friction and reduce the risk of fruit slippage during grasping. The insert was fixed to the gripper finger by adhesive bonding, allowing easy replacement or adjustment. This ribbed geometry provides a compliant interface that adapts to variations in fruit shape and surface irregularities, while maintaining a gentle but stable grip. The gripper insert was dimensioned to accommodate fruits with diameters ranging from 20 mm to 30 mm, ensuring reliable contact and sufficient compliance for both small and large specimens without causing mechanical damage.

Each gripper finger was actuated electrically, allowing a maximum fingertip force of 1,5 N with an average actuation time of 700 ms. However, due to the compliant TPU 85A metamaterial insert and the ribbed geometry of the soft finger, the effective contact force transmitted to the fruit surface was significantly attenuated and distributed over a larger area, ensuring safe manipulation of delicate produce. For fruits with diameters of 20–30 mm, and typical contact patches of ~60–126 mm² (e.g., ~60° wrap and 6–8 mm insert width), the resulting contact pressures are ≈12–24 kPa per finger, remaining within the imposed ≤30–40 kPa limit. Initial trials were therefore conducted at 15–25 kPa and increased only if grasp stability was insufficient.

Fig. 2 - a) 3D Model of the metamaterial insert; b) Assembled metamaterial insert on the robot's gripper

In the development of autonomous robotic systems for fruit harvesting, accurately determining whether a fruit has been successfully grasped by the gripper is a critical aspect that significantly impacts the efficiency, reliability, and quality of the entire process. This requirement becomes even more important in the context of soft grippers fabricated from mechanical metamaterials, which are designed to adapt their geometry during interaction with delicate, irregularly shaped objects such as fruits. The dynamic and compliant nature of these grippers introduces additional complexity, as the deformation of the gripper itself must be interpreted in real time to evaluate grasp success.

Thus, computer vision was employed to evaluate whether the fruit was correctly grasped, considering the shape deformation of the metamaterial gripper during interaction. A convolutional neural network (CNN) was trained on images collected during the experiments, with the aim of automatically classifying each grasp as correct or incorrect.

The illustrated convolutional neural network (CNN) (Figure 3) architecture was designed for binary classification of robotic grasp outcomes, specifically to determine whether an object has been successfully grasped ("Grasped Correct") or not ("Grasped Incorrect") (see Figure 5). The model takes as input an RGB image of dimensions 224×224×3, which is first processed through a convolutional layer with 32 filters of size

3×3, activated by a ReLU function. This is followed by a 2×2 max pooling operation, which reduces the spatial dimensionality and retains the most prominent features. The second convolutional block consists of 128 filters of size 3×3, again followed by a 2×2 max pooling layer, allowing the model to extract more abstract and complex visual features relevant to the classification task. The resulting feature maps are flattened into a one-dimensional vector and passed through a fully connected dense layer with 128 neurons and ReLU activation. To mitigate overfitting, a dropout layer with a rate of 0.5 is introduced before the final classification output. The network concludes with a binary output layer, which assigns a probability to one of the two classes, indicating whether the grasp attempt was correct or incorrect. This architecture balances simplicity and efficiency, making it well-suited for deployment on real-time robotic systems with limited computational resources.

The training configuration was meticulously designed to ensure both efficient convergence and robust generalization of the convolutional neural network (CNN) model developed for classifying robotic grasp outcomes. The model was trained to distinguish between three mutually exclusive classes—Grasped_Correct, Grasped_Failed, and No_Object-each representing a distinct outcome of the robotic grasping process. To support this multi-class classification task, the categorical cross-entropy loss function was employed. This loss is particularly well-suited for scenarios in which the model must output a probability distribution over discrete class labels, and it penalizes incorrect predictions proportionally to their confidence, thereby encouraging calibrated and accurate outputs.

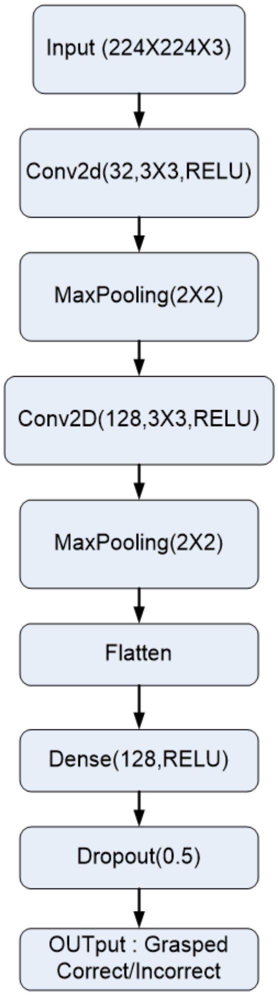

Fig. 3 – CNN Architecture

For optimization, the Adam optimizer was utilized with an initial learning rate set to 0.001, a widely accepted default that provides a reliable compromise between convergence speed and gradient stability. Adam's adaptive learning rate mechanism combines the advantages of both AdaGrad and RMSProp, enabling it to perform well in problems with sparse gradients and non-stationary objectives, which are typical in image-based classification tasks involving CNNs. The learning rate was kept constant during training; however, learning rate scheduling or decay strategies could be explored in future work to further refine performance.

Training was conducted over a range of 30 to 50 epochs, with the final number determined dynamically using an early stopping strategy. This regularization technique was employed to mitigate the risk of overfitting by monitoring the validation loss across epochs and terminating training automatically once no further improvement was observed after a defined patience threshold. This approach ensures that the model retains its ability to generalize well to unseen data while avoiding unnecessary computational expense. A batch size of 32 was selected after empirical tuning, as it offered a favourable balance between model convergence stability, memory efficiency, and training speed on standard GPU hardware. The chosen batch size also promoted smoother gradient updates, which are particularly beneficial for convergence in convolutional neural network (CNN) architectures trained on image data. Additionally, this configuration was found to provide an optimal trade-off between training accuracy and generalization performance across multiple experimental runs.

The testing of convolutional neural networks (CNNs) developed for detecting the grasp state of a fruit by a robotic gripper was conducted on a custom dataset containing images labelled according to three classes: Grasped_Correct, Grasped_Failed, and No_Object. The test set consisted of 150 images, proportionally balanced among the classes (50 images per class), sourced from controlled laboratory sessions using the 3D-printed soft gripper and fixed RGB camera, with a training/validation/test split of 70/20/10 %. Data augmentation included horizontal flips, and brightness adjustments of ± 20%. Model performance was evaluated using accuracy, precision, recall, and F1-score, with confusion matrices generated for each experimental run.

Experiments were conducted under controlled laboratory conditions at a temperature of 22 °C, using LED illumination at 300 lux, with a uniform background to minimize visual noise.

This methodological framework ensured a consistent validation of both the mechanical design and the vision-based algorithm, providing a basis for the performance analysis presented in the Results section.

RESULTS

The CNN model was trained on a Pc Intel I9, with NVIDIA® GeForce RTX™ 4090 graphic card and DDR5 128GB RAM. The software environment comprised Visual Studio, Cuda 11.8, Python 3.8 and OpenCV 14. As shown in table 1, the model correctly identified 94% of correct grasps, 91.6% of failed grasps, and 95.9% of cases where no object was present between the fingers. Precision for correct grasps was ~96.2%, while recall reached 94%, indicating that the model correctly identified most true positives but occasionally misclassified failed grasps.

Table 1

Performance metrics

Class	Precision [%]	Recall [%]	F1-score [%]
Correct grasp	96.2	94	95.1
Failed grasp	93.8	91.6	92.7
No object	91	95.9	93.4

Table 2 highlights that there was a tendency for the model to sometimes confuse failed grasps with correct ones, which can be explained by partial visual similarity (e.g., unilateral contact with the fruit, but not centred). The confusion matrix highlighted that 3% of failed grasps were misclassified as correct, mainly due to unilateral fruit contact or partial occlusion.

Table 2

Confusion matrix

Predicted / Actual	Correct grasp	Failed grasp	No object
Correct grasp	94%	3%	2%
Failed grasp	3%	91.50%	2%
No object	3%	5.50%	94%

With an average processing time below 100 ms per image, the system can operate in real time, supporting cycle times compatible with robotic harvesting tasks.

Table 3

Mean inference time

Metric	Value [ms]
Mean inference time	87
Standard deviation	6
Min / Max	75 / 98

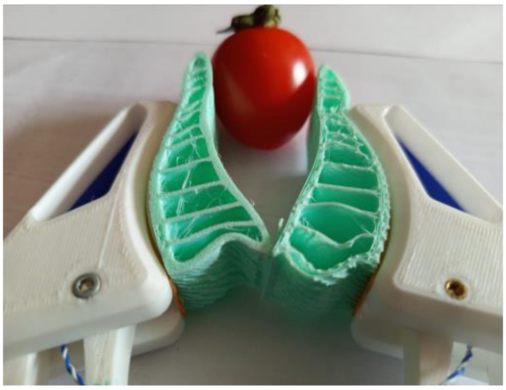

Fig. 4 - Example of incorrect grasp training image

Based on visual analysis of the model's predictions, some relevant aspects were observed. Thus, the model is sensitive to partial occlusions, especially when one finger of the gripper partially covers the tomato. This often causes the system to misinterpret the presence or position of the object, particularly when the gripper is in motion or when the occlusion overlaps the central region of the fruit.

Non-uniform lighting in the scene, especially when light is concentrated or shadowed on one side of the frame, significantly affects the network's ability to distinguish between Grasped_Failed and Grasped_Correct. Performance decreased under non-uniform lighting, where accuracy dropped by 5%, highlighting the model's sensitivity to illumination changes. Figure 4 shows an example of image trained where grasp failed. This highlights the model's reliance on consistent illumination and suggests a need for either data augmentation addressing lighting variation or adaptive preprocessing techniques. Off-centre grasps, where the tomato is held closer to one finger than the other, are often mistakenly classified as Grasped_Correct. Although visually similar to a proper grasp, these cases may indicate a fragile or unstable grip in practice. This reveals the model's current limitation in interpreting spatial symmetry as a critical grasp success factor. To address these issues, future iterations of the model could benefit from incorporating spatial attention mechanisms, or even integrating depth information to better resolve occluded or asymmetric cases. Moreover, training on a more diverse set of lighting conditions and grasp variations could enhance the model's robustness in real-world applications.

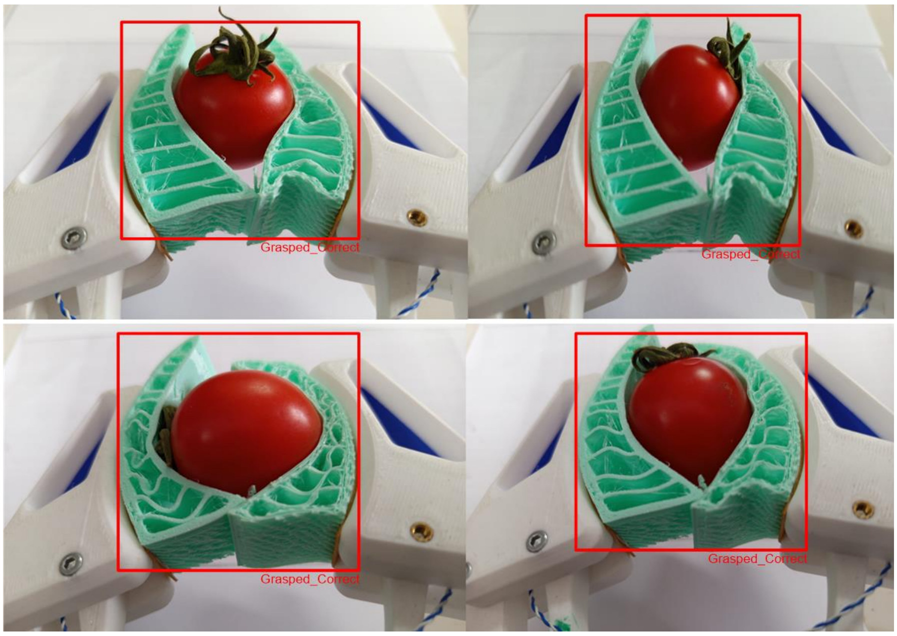

Fig. 5 – Identification results for correct grasped tomatoes

CONCLUSIONS

This study presented the design and experimental validation of a vision-based method for detecting grasp success in fruit harvesting using a soft gripper fabricated from mechanical metamaterials. The developed system integrates computer vision with a convolutional neural network (CNN) to classify the quality of the grasp based on the visual deformation of the gripper during interaction with fruits.

The study demonstrated that the proposed computer vision approach can reliably determine the success of fruit grasping when using a metamaterial soft gripper. The system achieved an accuracy of 94.0% for correct grasps, 91.5% for failed grasps, and 95.9% for no-object cases. The average inference time was 87 ms, with a variation between 75 and 98 ms, which ensures real-time applicability in harvesting tasks. These results confirm that vision-based monitoring of gripper deformation provides robust feedback for adaptive control strategies, enabling safe and efficient harvesting of delicate fruits.

The proposed approach improves harvesting efficiency by allowing the robot to detect and correct failed attempts, reducing cycle times and minimizing potential damage to fragile produce. By exploiting the deformation characteristics of metamaterial-based grippers, the method enables adaptive and safe manipulation of fruits with varying ripeness, geometry, and texture.

However, the experiments were carried out under controlled laboratory conditions, which may not fully reflect the challenges encountered in real greenhouse environments, such as variable lighting, occlusions, and dynamic plant structures. Future research will therefore focus on testing the system in commercial greenhouse settings, expanding the training dataset to improve robustness, and integrating multimodal feedback by combining vision with tactile or force sensing to further enhance reliability and safety.

ACKNOWLEDGEMENT

This work was supported by a grant of the Ministry of Agriculture and Rural Development, contract of sectorial financing, ADER 2023-2026 type, no. 25.1.1, “Technology for robotized harvesting of *Solanaceae* family vegetables in greenhouses and solariums, using artificial intelligence”.

REFERENCES

[1] Bharad, N.B., Khanpara, B.M. (2024). Agricultural fruit harvesting robot: An overview of digital agriculture, *Plant Archives*, 24 (Special Issue GABELS), pp. 154–160, DOI: https://doi.org/10.51470/PLANTARCHIVES.2024.v24.SP-GABELS.023.

[2] Elfferich, J.F., Dodou, D., Della Santina, C. (2022). Soft robotic grippers for crop handling or harvesting: A review, *IEEE Access*, 10, pp. 75428–75443, DOI: https://doi.org/10.1109/ACCESS.2022.3190863.

[3] Ge, Y., Xiong, Y., Tenorio, G.L., From, P.J. (2019). Fruit localization and environment perception for strawberry harvesting robots, *IEEE Access*, 7, pp. 147642–147652, DOI: https://doi.org/10.1109/ACCESS.2019.2946369.

[4] Grady, P., Collins, J.A., Brahmbhatt, S., Twigg, C.D., Tang, C., Hays, J., Kemp, C.C. (2022). Visual pressure estimation and control for soft robotic grippers, *Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3628–3635, IEEE, DOI: https://doi.org/10.1109/IROS47612.2022.9981241.

[5] He, Z., Li, Z., Ding, X., Li, K., Shi, Y., Cui, Y. (2023). Design and experiment of end effect for kiwifruit harvesting based on optimal picking parameters. *INMATEH – Agricultural Engineering, 69*(1), pp. 325–334. DOI: https://doi.org/10.35633/inmateh-69-30.

[6] Hegde, C., Mysa, R.C., Chooi, A., Dontu, S., Tan, J.M.R., Wong, L.H., Valdivia y Alvarado, P., Magdassi, S. (2024). 3D-printed mechano-optic force sensor for soft robotic gripper enabled by programmable structural metamaterials, *Advanced Intelligent Systems*, 6(9), art. no. 202400057, DOI: https://doi.org/10.1002/aisy.202400057.

[7] Huang, X., Guo, W., Liu, S., Li, Y., Qiu, Y. (2022). Flexible mechanical metamaterials enabled electronic skin for real-time detection of unstable grasping in robotic manipulation, *Advanced Functional Materials*, 32(44), 2109109, DOI: https://doi.org/10.1002/adfm.202109109.

[8] Kaur, M., Kim, W.S. (2019). Toward a smart compliant robotic gripper equipped with 3D-designed cellular fingers, *Advanced Intelligent Systems*, 1(7), 1900019, DOI: https://doi.org/10.1002/aisy.201900019.

[9] Li, Z., Yuan, X., Wang, C. (2022). A review on structural development and recognition–localization methods for end-effector of fruit–vegetable picking robots, *International Journal of Advanced Robotic Systems*, 19(4), DOI: https://doi.org/10.1177/17298806221104906.

[10] Liu, J., Liu, Z. (2024). The vision-based target recognition, localization, and control for harvesting robots: A review, *International Journal of Precision Engineering and Manufacturing*, 25(2), pp. 409–428, DOI: https://doi.org/10.1007/s12541-023-00911-7.

[11] Matache, M.G., Găgeanu, I., Brăcăcescu, C., Cristea, O.D., Cristea, R.D. (2025). Robotic harvesting methods for vegetables of the Solanaceae family in greenhouses and solariums using specialized grippers – a review, *Acta Horticulturae*, 1433, pp. 323–330, DOI: https://doi.org/10.17660/ActaHortic.2025.1433.41.

[12] Mohammadi, A., Hajizadeh, E., Tan, Y., Choong, P., Oetomo, D. (2023). A bioinspired 3D-printable flexure joint with cellular mechanical metamaterial architecture for soft robotic hands, *International Journal of Bioprinting*, 9(3), 696, DOI: https://doi.org/10.18063/ijb.696.

[13] Navas, E., Shamshiri, R. R., Dworak, V., Weltzien, C., Fernández, R. (2024). Soft gripper for small fruits harvesting and pick and place operations, Frontiers in Robotics and AI, 10, p. 1-14, DOI: https://doi.org/10.3389/frobt.2023.1330496.

[14] Nguyen, V.P., Dhyan, S.B., Mai, V., Han, B.S., Chow, W.T. (2023). Bioinspiration and biomimetic art in robotic grippers, *Micromachines*, 14(9), 1772, DOI: https://doi.org/10.3390/mi14091772.

[15] Pyo, S., Park, K. (2024). Mechanical metamaterials for sensor and actuator applications, *International Journal of Precision Engineering and Manufacturing – Green Technology*, 11, pp. 291–320, DOI: https://doi.org/10.1007/s40684-023-00549-w.

[16] Seo, D., Oh, I.-S. (2025). Gripping Success Metric for Robotic Fruit Harvesting, *Sensors*, 25(1), p.181, DOI: https://doi.org/10.3390/s25010181.

[17] Sui, D., Zhu, Y., Zhao, S., Wang, T., Agrawal, S.K., Zhang, H., Zhao, J. (2022). A bioinspired soft swallowing gripper for universal adaptable grasping, *Soft Robotics*, 9(1), pp. 36–56, DOI: https://doi.org/10.1089/soro.2019.0106.

[18] Shintake, J., Cacucciolo, V., Floreano, D., Shea, H. (2018). Soft robotic grippers, *Advanced Materials*, 30(29), 1707035, DOI: https://doi.org/10.1002/adma.201707035.

[19] Su, M., Guan, Y., Huang, D., Zhu, H. (2021). Modeling and analysis of a passively adaptive soft gripper with the bio-inspired compliant mechanism, *Bioinspiration & Biomimetics*, 16(5), 055005, DOI: https://doi.org/10.1088/1748-3190/ac07f7.

[20] Tawk, C., Mutlu, R., Alici, G. (2022). A 3D printed modular soft gripper integrated with metamaterials for conformal grasping, *Frontiers in Robotics and AI*, 8, 799230, DOI: https://doi.org/10.3389/frobt.2021.799230.

[21] Tang, Y., Chen, M., Wang, C., Luo, L., Li, J., Lian, G., Zou, X. (2020). Recognition and localization methods for vision-based fruit picking robots: A review, *Frontiers in Plant Science*, 11, 510, DOI: https://doi.org/10.3389/fpls.2020.00510.

[22] Visentin, F., Castellini, F., Muradore, R. (2023). A soft, sensorized gripper for delicate harvesting of small fruits, *Computers and Electronics in Agriculture*, 213, 108202, DOI: https://doi.org/10.1016/j.compag.2023.108202.

[23] Wang, D., Dong, L., Gu, G. (2022). 3D printed fractal metamaterials with tunable mechanical properties and shape reconfiguration, *Advanced Functional Materials,* 32(44), 2208849, DOI: https://doi.org/10.1002/adfm.202208849.

[24] Xie, B., Jin, M., Duan, J., Li, Z., Wang, W., Qu, M., Yang, Z.. (2024). Design of Adaptive Grippers for Fruit-Picking Robots Considering Contact Behavior, *Agriculture*, 14(7), 1082, DOI: https://doi.org/10.3390/agriculture14071082.

[25] Zhou, J., Chen, S., Wang, Z. (2017). A soft-robotic gripper with enhanced object adaptation and grasping reliability, *IEEE Robotics and Automation Letters*, 2(4), pp. 2287–2293, DOI: https://doi.org/10.1109/LRA.2017.2716445.