Precision Engineering 88 (2024) 367–381

Contents lists available at ScienceDirect

Precision Engineering

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

A novel vision-based multi-functional sensor for normality and position measurements in precise robotic manufacturing☆

Mohamad Halwani [a,*,1], Abdulla Ayyad [a,1], Laith AbuAssi [a], Yusra Abdulrahman [a,b], Fahad Almaskari [b], Hany Hassanin [c], Abdulqader Abusafieh [d], Yahya Zweiri [a,b]

[a] Advanced Research and Innovation Center (ARIC), Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates
[b] Department of Aerospace Engineering, Khalifa University, Abu Dhabi, United Arab Emirates
[c] School of Engineering, Technology, and Design, Canterbury Christ Church University, Canterbury CT1 1QU, UK
[d] Research and Development, Strata Manufacturing PJSC, Al Ain, United Arab Emirates

ARTICLE INFO

Keywords:
Multi-functional sensor
Vision-based tactile sensing
Precise manufacturing
Robot deburring

ABSTRACT

Cobots play an essential role in the fourth industrial revolution and the automation of complex manufacturing processes. However, cobots still face challenges in achieving high precision, which obstructs their usage in precise applications such as the aerospace industry. Nonetheless, advances in perception systems unlock new cobot manufacturing capabilities. This paper presents a novel multi-functional sensor that combines visual and tactile feedback using a single optical sensor, featuring a moving gate mechanism. This work also marks the first integration of Vision-Based Tactile Sensing (VBTS) into a robotic machining end-effector. The sensor provides vision-based tactile perception capabilities for precise normality control and exteroceptive perception for robot localization and positioning. Its performance is experimentally demonstrated in a precise robotic deburring application, where the sensor achieves the high-precision requirements of the aerospace industry with a mean normality error of 0.13° and a mean positioning error of 0.2 mm. These results open a new paradigm for using vision-based sensing for precise robotic manufacturing, which surpasses conventional approaches in terms of precision, weight, size, and cost-effectiveness.

1. Introduction

"Factory of the future" is a term introduced in the 1980s [1] to describe an evolving era led by robotics, automation, and artificial intelligence. For instance, the fourth industrial revolution, or Industry 4.0, spotlights industrial automation and perceptive robotic systems that can perform various tasks like: grasping [2], packaging [3], as well as precise machining and manufacturing processes in unstructured environments [4]. Robots provide high precision, consistency, and operational efficiency compared to manual activity, making them valuable for automating precise manufacturing processes. Thereby, robotic automation boosts performance and production output with lower costs and error incidence [4–8].

The aforementioned advantages were key factors to the wide adoption of robots in several manufacturing sectors, including the automotive industry, whereas the aerospace industry is lagging in employing robotic solutions due to the higher level of precision needed. In the aerospace industry, thousands of holes are drilled in aircraft structures to act as mechanical connection points [9,10]. The positional accuracy of the drilled holes must be within 0.5 mm to be considered acceptable [11,12]. Also, the perpendicularity error of drilled holes must not exceed 2° to ensure that the rivet connections work as expected [13–15]. Failure to meet aerospace standards for drilled holes has a negative effect on the connection quality and could reduce fatigue life due to cyclic loading below the material strength [16]. The percentage of fatigue failure accidents of the aircraft body due to the joints of the structure is reported to be 70%, while rivet holes cause 80% of the fatigue cracks [17]. Furthermore, a drilling burr can be formed if the machining tool is not precisely aligned with the workpiece. This results in poor surface quality, assembly problems, and reduced fatigue life of the structure [18]. Perpendicularity and position errors are two of the main geometrical tolerances of a drilled hole, and obtaining precise

☆ This paper was recommended by Associate editor Prof. R. Leach.

This work was supported by the Advanced Research and Innovation Centre (ARIC), which is jointly funded by STRATA Manufacturing PJSC (a Mubadala company), Khalifa University of Science and Technology, in part by Sandooq Al Watan under Grant SWARD-S22-015.

* Corresponding author.
E-mail address: mohamad.halwani@ku.ac.ae (M. Halwani).

[1] Equal contribution to this work.

https://doi.org/10.1016/j.precisioneng.2024.02.015
Received 27 May 2023; Received in revised form 7 November 2023; Accepted 24 February 2024
Available online 26 February 2024

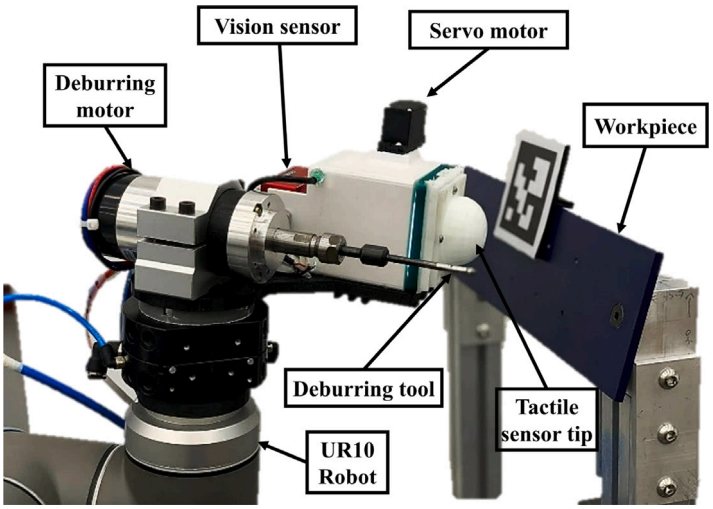

Fig. 1. The proposed system for combined normality and position measurement and control. A multi-functional vision-based sensor can be used for precise alignment and guidance of the robotic end-effector utilizing only a single vision sensor.

measurements of the above quantities is fundamental to achieving those tolerances.

These challenges become more apparent when utilizing cobots, which, despite being safer, more cost-effective and easier to integrate into the factory, provide lower precision than traditional industrial robots. This is especially true when the robot is operating in unstructured environments with limited information about its surroundings and target workpieces. As a result, enhancing the precision of cobots has been an active area of research, where works in literature can be divided into two approaches: the first approach aims to augment the cobot's precision through more advanced low-level control of the robot's joints. In contrast, the second approach focuses on using sensory feedback and closed-loop control to improve the robot's repeatability. The current advancements in sensor technology have been a leading factor to significant growth in the adoption of robots in automated manufacturing [4]. Sensors have become an essential aspect in the development of intelligent robotic manufacturing systems as they allow for real-time monitoring and control of the manufacturing process, resulting in improved system performance, efficiency, and safety [4, 19].

Visual and tactile perception is crucial for cobot operations in real-world and industrial environments, where they provide useful feedback that can be utilized to improve the precision and repeatability of cobots, as well as increase their applicability and safety. With the advancement of these technologies, cobots will be able to execute dexterous manipulation tasks that will serve the manufacturing industry and boost productivity, efficiency, and workplace safety. In our previous work [20], we addressed the positional accuracy of robotic machining using neuromorphic vision-based control and achieved an average positional error of less than 0.1 mm. In this work, we address the normality/perpendicularity errors of robotic machining and introduce a novel vision-based sensor capable of measuring both the tool's normality and position with tolerances below 1° and 0.2 mm, respectively. We also present the control law that uses the sensor measurements to precisely position and align the machining end-effector relative to an arbitrarily placed workpiece. The proposed vision-based sensor mounted on a collaborative robot can be seen in Fig. 1.

In recent years, different methods have been proposed in the literature to measure and align a robotic end-effector with the normal axis of a workpiece. Several studies employ an array of distance sensors to adjust the end-effector's normality; but use different algorithms, arrangements, and calibration methods [11,12,19,21–24]. Other well-studied methods include utilizing 3D laser scanners [25,26], fringe projection sensors [27], or force sensors [28] for normality measurements. Despite the high precision of the aforementioned methods, they have limitations including high cost, complex calculations, excessive instrumentation requirements, relatively high weight, and limited functionality. Almost all of the aforementioned sensors only provide measurements on the tool's normality and require additional sensors to be used for positioning. For instance, the work in [12] uses a set of four contact-less distance sensors to measure the normality of a workpiece, while a 2D camera and three laser sensors are used to control the robot's positioning. The multi-sensor approach not only increases the payload of the robot but also increases the development and maintenance costs. Furthermore, it increases system complexity, requiring advanced calibration and data processing. In this paper, we propose a novel vision-based tactile sensor that uses a single camera for both robot positioning and normal alignment. Our novel design combines both perception capabilities in a single modality, while prior works in the literature address positioning and normality as separate problems with an independent sub-system for each. Consequently, our sensor significantly outperforms conventional options by offering greater functionality while reducing cost, weight, and the need for complex instrumentation and wiring.

In contrast to laser projection techniques, Vision-Based Tactile Sensors (VBTS) offer a more robust and adaptable option for precise measurements. Unlike laser systems, which often come with high costs, difficult computations, and excessive instrumentation requirements, VBTS simplifies the system by reducing both complexity and cost. They also excel in flexibility, easily adapting to a variety of surface textures, materials, and shapes without the need for special surface treatments like laser projection techniques. Moreover, VBTS are generally considered safer, as they do not involve the hazards associated with laser use. With the potential to serve as a multi-purpose sensor, VBTS are capable of sensing other metrics such as vibrations and contact forces. Vision-based tactile sensors consist of optical sensors that monitor the interior surface of a low-cost elastic body (referred to as the tactile surface). Internal markers that re-arrange due to the elastic body deformation upon any physical contact contain rich tactile information that can be decoded to estimate contact forces, normality, and various measurements from the contact surface, as illustrated in Fig. 2. Vision-based tactile sensors have gained interest in the robotics community because they offer advantages over conventional tactile sensors. These advantages include high spatial resolution, design simplicity, and low instrumentation requirements [29]. Studies have shown the benefits of using vision-based tactile sensors in applications such as object manipulation [30–32], slip detection [33–35], contact

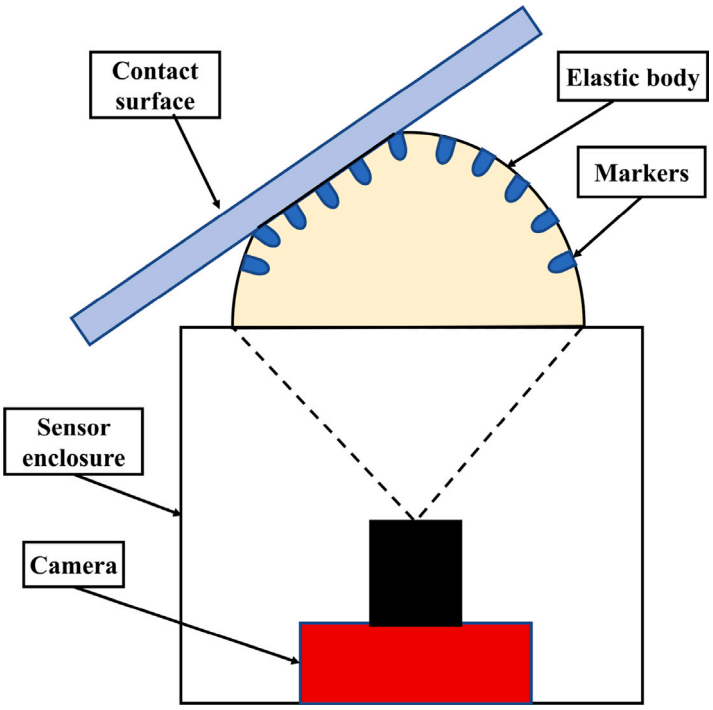

Fig. 2. The configuration and principle of operation of vision-based tactile sensors. A camera observes the deformation of an elastic body once it presses against a contact surface. The camera's observations contain rich information from the markers' rearrangement due to the elastic body deformation, which can be used to estimate the contact pose in terms of normality and depth.

forces measurement [31,36,37], contact pose estimation [38,39], and texture recognition [33,40]. In [41], stereo cameras for vision-based tactile sensor were used to estimate force distribution, contact position, angular displacement, and contact area. However, they reported an angle error of 2.235°, which is unsatisfactory for machining in the aircraft industry. In [38], a vision-based tactile sensor was trained to predict the 3D pose of objects and edges that are in contact to achieve object manipulation. The study achieved a Mean Absolute Error (MAE) of 0.3° for roll and pitch measurements. Another study [42] measured the contact pose components using tactile sensors, but for a grasping task, they achieved MAE of 0.33° and 0.28° for roll and pitch, respectively. However, no studies to date have considered the application of vision-based tactile sensors in precise robotic machining, nor do they provide the required tolerances or controls required for manufacturing processes in the aerospace industry. Additionally, prior works limit the use of the imaging sensor for proprioceptive sensing to provide a specific tactile sensing task and do not exploit the flexibility of imaging sensors for additional exteroceptive tasks. Our work overcomes this by using a single camera for both proprioceptive and exteroceptive sensing. Table 1 compares several recent studies on robotic perception technologies in terms of functionality and application. It shows that this paper constitutes the first work in literature to combine proprioceptive, tactile sensing and exteroceptive sensing for positioning using a single imaging sensor. Although prior work on VBTS has developed algorithms to estimate tactile quantities such as: contact pose, slippage, and force; none have ever investigated using the same camera of VBTS for exteroceptive perception functions. This idea significantly simplifies the system, reducing both complexity and cost. While a dual-camera setup might seem simpler, it would increase the cost and complexity for high-speed industrial applications, where robust, high-frame-rate cameras are often required. In addition, this added feature of our design facilitates wider adoptions and applications of VBTS, especially since most robot manipulation tasks require a combination of vision and touch sensing. Furthermore, this paper presents the first vision-based tactile sensor for precise robotic machining, while prior work on VBTS mostly targets challenges in gripping and grasping.

In this paper, we propose a novel vision-based tactile sensor for the normality and positional control of precise robotic machining tasks. The sensor combines proprioceptive and exteroceptive capabilities, where a single camera observes both the surrounding environment for robot guidance and the inner surface of an elastic body to estimate tactile properties [47]. The sensor features a moving gate mechanism, which enables switching between exteroceptive perception for robot localization and proprioceptive sensing for tactile feedback. The sensor enables the robot to localize a workpiece in 6-DoF, and uses position-based visual servoing (PBVS) to control the robot's position accordingly. As the robot achieves contact with the workpiece, the sensor's camera observes the deformation of the tactile surface, and utilizes Convolutional Neural Networks (CNNs) to relate the obtained image of surface deformation to the required contact properties in terms of normality and depth. Based on the sensor's measurement, the robot regulates the machining tool's pose as per the tolerances required in the aerospace industry. We validated the sensor presented experimentally in a robotic deburring application and presented both quantitative and qualitative analyses of the obtained results. A video demonstration of the sensor and corresponding control laws in action can be accessed through this link: https://drive.google.com/file/d/1-JCEoXYHVApuky3Aln8KjX2k4ZXQiSRy/view?usp=sharing. The main contributions of this paper can be summarized as follows:

- We develop a novel multi-functional sensor configuration for robotic manufacturing that combines proprioceptive and exteroceptive sensing using only a single camera. This proposed design offers improvements in cost, weight, and design simplicity over prior work in the literature, pioneering the application of tactile sensing in robotic manufacturing tasks.
- For the first time, we propose vision-based tactile sensing for normality adjustment in precise robotic machining. We develop the corresponding perception and control algorithms, which employ a Convolutional-Neural-Network architecture and training methodology. This approach outperforms State-of-The-Art, achieving a mean absolute error of 0.13° and significantly enhancing precision in robotic machining.

Table 1
An overview of the scope and approach of recent works on robotic perception technologies. This paper is the first work in literature to address both proprioceptive and exteroceptive perception requirements using a signal optical sensor, and constitutes the first VBTS sensor for precise machining applications.

Study	Proprioceptive perception	Exteroceptive perception	Sensors	Application
E. Pso. et al. [42]	Yes	No	VBTS	Grasping
N. Lepora et al. [38]	Yes	No	VBTS	Contact estimation
R. Muthusamy et al. [43]	Yes	No	VBTS	Grasping
A. Gupta et al. [44]	Yes	No	VBTS	Contact estimation
M. Lambeta et al. [30]	Yes	No	VBTS	In-hand manipulation
A. Rigi et al. [35]	Yes	No	Event-based camera	Slip detection
T. Olsson et al. [28]	Yes	No	Force sensor	Robotic drilling
B. Mei et al. [45]	No	Yes	2D Camera and laser sensors	Robotic drilling
T. Jiang et al. [46]	No	Yes	Multiple cameras	peg-in-hole
T. Song et al. [26]	No	Yes	Laser sensors	Surface normal measurement
G. Rao et al. [27]	No	Yes	2D camera and fringe projection sensor	Robotic drilling
P. Yuan et al. [11]	No	Yes	2D camera and laser sensors	Robotic drilling
A. Ayyad et al. [20]	No	Yes	Neuromorphic vision	Robotic drilling
This work	**Yes**	**Yes**	**VBTS**	**Robotic deburring**

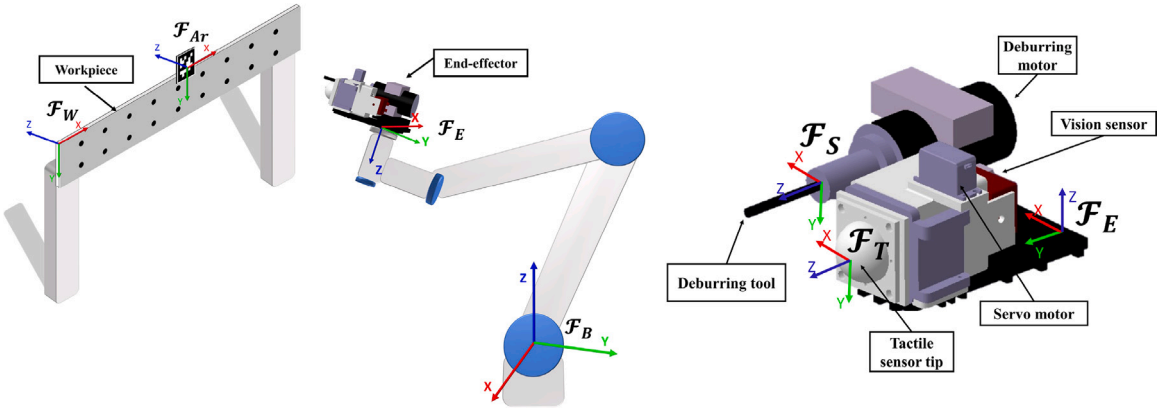

Fig. 3. Robotic deburring setup that illustrates the coordinate frames used for navigation and control.

- We formulate and validate the visual guidance and control laws that utilize the camera's exteroceptive feedback to precisely position the robotic tool relative to a target workpiece with 0.2 mm accuracy.
- We perform experimental validation and testing of the proposed sensor in a robotic deburring setup and present a detailed qualitative and quantitative analysis of the obtained results. Results show that the proposed sensor and perception algorithms provide the positioning and normality requirements of automated manufacturing in the aerospace industry.

The rest of this paper is organized as follows: Section 2 outlines the setup and configuration of the proposed deburring robotic setup. Section 3 describes the working principle and the methodology. In Section 4, we present both qualitative and quantitative results of the proposed system. Finally, in Section 5, we conclude the paper.

2. Robotic setup

The overall setup of the multi-functional vision-based sensor with the robotic manipulator can be seen in Figs. 1 and 3. The presented system uses the camera feedback to both position and align a deburring tool with pre-drilled holes in a workpiece. We used Universal Robot's UR10,[2] which provides a repeatability value of 0.1 mm. To address the full reach requirement for larger workpieces, the robot can be mounted on a mobile platform, enabling it to navigate and adapt to large-scale aircraft components, as demonstrated in our prior work in [20]. Fig. 4 shows the end-effector configuration, which consists of the deburring-tool and our multi-functional vision-based sensor. Integrating the sensor into the robot system is a straightforward process, where the sensor can be easily mounted onto the robot's end-effector, allowing for seamless integration without the need for extensive modifications or instrumentation requirements. The images of the DAVIS346 camera[3] with a resolution of 346 × 260 were used for perception. The camera integrated into the system is capable of providing both images and events, making it versatile for perception tasks. In the context of tactile sensing, both image-based and event-based approaches have been utilized in literature [40,48]. An upcoming study will specifically investigate and compare the effectiveness of image-based and event-based Visual Tactile Sensing (VBTS) for robotic machining applications. It is important to note that the design and algorithms reported in this paper are not limited to a specific camera model, such as the DAVIS346. Instead, any RGB camera can be employed, offering flexibility in sensor selection and potentially reducing size and weight according to specific requirements. For the purpose of robot navigation and control, we define the following frames of reference:

- $\mathcal{F}_B$: The robot base coordinate frame.
- $\mathcal{F}_E$: The robot end-effector coordinate frame.
- $\mathcal{F}_S$: The deburring tool coordinate frame.
- $\mathcal{F}_T$: The tactile surface coordinate frame.
- $\mathcal{F}_C$: The camera coordinate frame.
- $\mathcal{F}_W$: The workpiece coordinate frame.

[2] UR10: https://www.universal-robots.com/media/50880/ur10_bz.pdf.

[3] DAVIS346: https://inivation.com/wp-content/uploads/2019/08/DAVIS346.pdf.

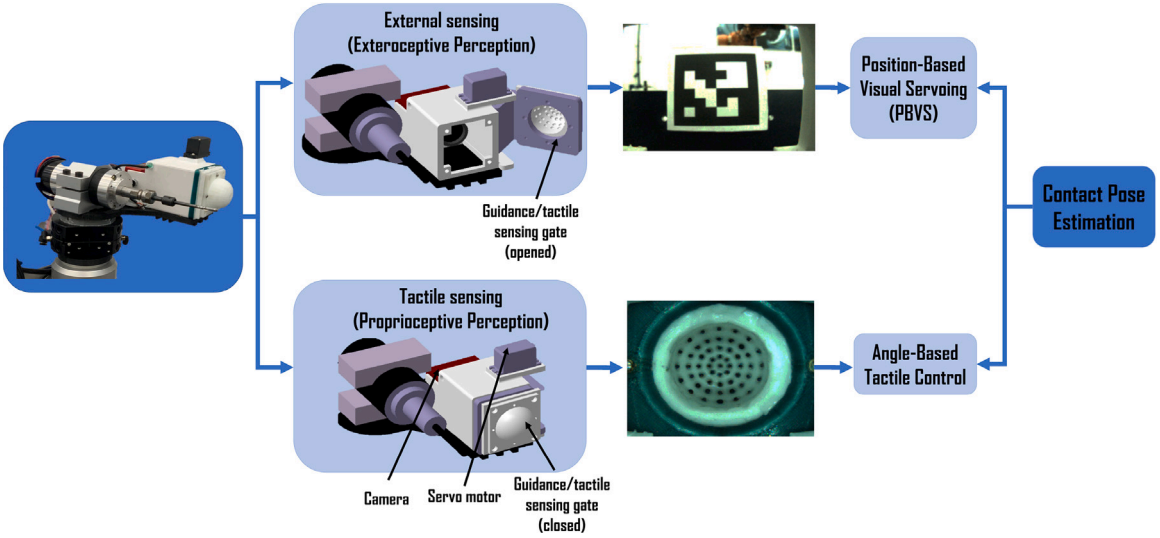

Fig. 4. The end-effector configuration shows the multi-functional vision-based sensor with the deburring tool. When the sensor gate is opened, the output of the vision sensor shows the surrounding environment for the robot guidance task (Exteroceptive perception). However, while the sensor gate is closed, the vision sensor output shows the sensor's internal surface for the tactile sensing task (Proprioceptive perception). Both of these tasks are executed using position-based visual servoing (PBVS) for robot positioning and angle-based tactile servoing (ABTS) for normality adjustment.

- $\mathcal{F}_A$: The fiducial/ArUco coordinate frame.
- $\mathcal{F}_{h^i}$: The coordinate frame of the ith reference hole.

We define the affine transformation matrix ${}_T T_S \in \mathbb{R}^{4\times4}$ that maps from coordinate frame $\mathcal{F}_S$ to $\mathcal{F}_T$ as follows:

$$ {}_T T_S = \begin{bmatrix} {}_T R_S & {}^T_T\vec{P}_s \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{1} $$

where, ${}_T R_S \in \mathbb{R}^{3\times3}$ is the rotation matrix, and ${}^T_a\vec{P}_b \in \mathbb{R}^3$ gives the relative position of point b to point a described in coordinate frame $\mathcal{F}_T$.

Also, solving the robot's forward kinematics allows calculating the transformation from $\mathcal{F}_E$ to $\mathcal{F}_B$ as follows:

$$ {}_B T_E = g(\theta), \qquad \theta \in \mathbb{C} \tag{2} $$

where $\mathbb{C}$ is the robot's configuration space, θ is the measured joint angles for the articulated manipulator, and $g(\theta)$ is a non-linear function that defines robot kinematics. Furthermore, the constants ${}_E T_C$, ${}_E T_T$, and ${}_c T_S$ may be determined using the geometrical calibration method detailed in [20,49]. As a result, ${}_B T_E$ and the calibrated transformations can be combined to compute ${}_B T_C$, ${}_B T_T$, and ${}_B T_S$.

The robot has to identify ${}_B T_{h^i}$ to perform the machining operation. We solve this transformation in two steps. In the first step, we estimate the hole's position ${}^B_B\vec{P}_{h^i}$ using exteroceptive visual guidance as explained in Section 3.5. We then estimate the hole's normal vector using visual tactile sensing measurements as explained in Section 3.4. Using these measurements, our visual and tactile control law precisely aligns the deburring tool with the target holes in the workpiece, as described in detail in Section 3.6.

3. Methods

3.1. Vision-based tactile sensing operation principle

Vision-based tactile sensors (VBTS) are a type of tactile sensor that uses cameras to detect changes in the internal structure of a low-cost elastic tactile interface resulting from the application of forces. The inner surface of the tactile interface often includes a collection of visual markers as shown in Fig. 2. When the sensor is pressed against an object, the tactile interface is deformed, causing the internal markers to rearrange in response to the magnitude and direction of the contact force as can be seen in Fig. 5. This allows the VBTS to detect and measure the forces and positions of objects in contact with the sensor.

Table 2
Design features of the tactile sensor elastic material and internal markers.

Parameter	Value
Elastomer material	rubber (Hardness: Shore 00–30)
Elastomer size	40 mm diameter
Markers material	plastic
No. of markers	53
Size of markers	2.5 mm
Size of the sensor enclosure (L × H × W)	20 cm × 12 cm ×10 cm
Sensor 3D printedenclosure material	ABS
Sensor overall weight	0.47 kg

VBTS offers advantages over conventional tactile sensors in terms of design simplicity and high spatial resolution [29]. Furthermore, since the camera is the only active component in the sensor, VBTS has considerably lower instrumentation requirements compared to other strain-gauge or piezo-electric tactile sensors. Despite those advantages, one of the main challenges of VBTS is the lifetime of the elastic tactile interface, which might suffer from permanent deformation over long-term use. However, as the tactile interface is made from low-cost silicon material, it can be easily replaced once its lifetime has passed. Another challenge with VBTS is the need for the development of novel specialized algorithms to infer tactile measurements from images of the deformation, which heavily relies on the precise detection and labeling of the visual markers in each image frame to accurately track their displacement. Several studies measured contact forces and position by tracking the markers' displacement directly when an external force is applied to the sensor's surface [50]. This can be time-consuming and lead to inaccurate tactile sensing measurements if certain markers are not detected [51]. To address this problem, we have used convolutional neural networks (CNNs) to extract useful information directly from visual feedback to estimate the contact pose in terms of normality and depth. This approach can improve the accuracy and efficiency of VBTS by eliminating the need for manual marker labeling.

3.2. Sensor design

The proposed multi-functional vision-based sensor comprises a 3D-printed housing structure that holds the camera and the sensor tip,

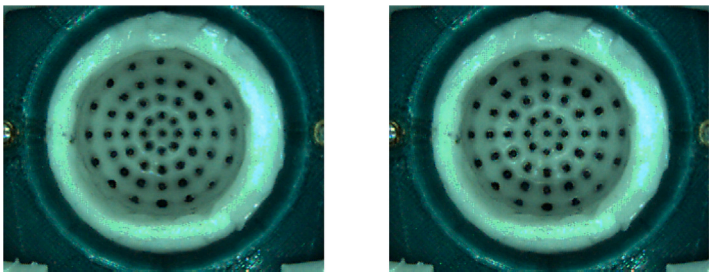

Fig. 5. The camera feedback shows the tactile interface: (a) before contact with the workpiece, and (b) after contact with the workpiece. The deformation of the tactile interface changes the arrangement pattern of the visual markers in the image. This pattern can then be used to infer different contact quantities including force, normality, and depth.

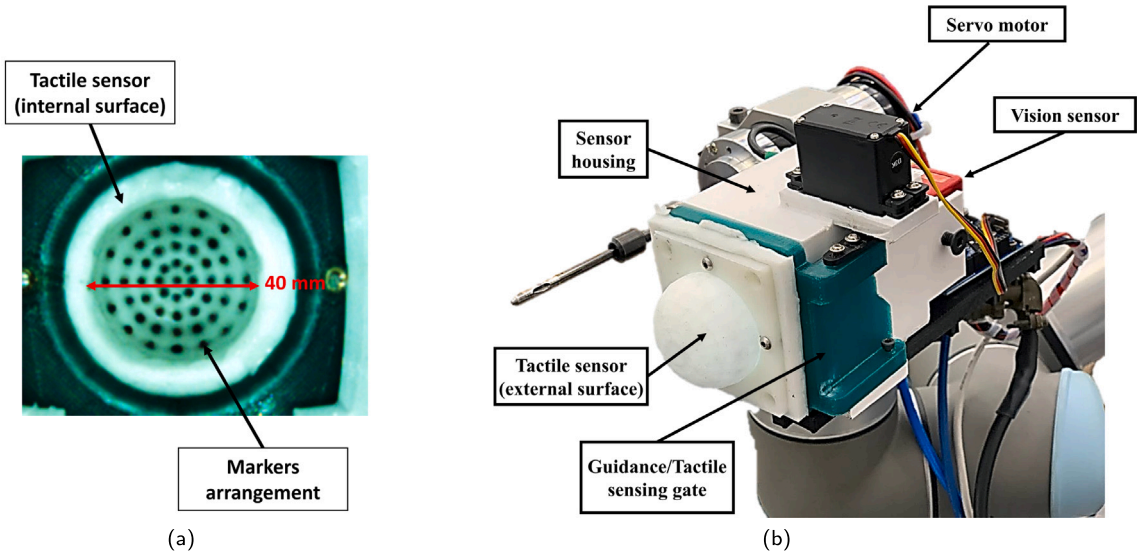

Fig. 6. (a) The internal surface of the tactile sensor shows the markers' arrangement. We chose the colors for the tactile sensor (white) and the 2 mm diameter beads (black) to have maximum color contrast for better visibility. (b) 3D-printed sensor enclosure compromising the elastic body mounted on a servo motor-controlled gate for exteroceptive/proprioceptive perception.

as shown in Fig. 6. The sensor tip is directly in front of the camera, with air as the sole medium between them. The sensor tip is a hollow hemispherical-shaped elastic surface fixed to the front side of a moving gate controlled by a servo motor. The hemispherical shape of the sensor tip helps evenly distribute the stress across the sensor and allows maximum displacement of markers, improving the tactile perception of the sensor [52]. The tactile surface has markers arranged on the internal surface and is made by molding a low-cost silicone soft material of shore hardness 00–30. The tactile surface elastomer material is chosen to have large deformations without breaking or losing its shape, as it allows the sensor to deform easily in response to tactile stimuli. The sensor design specifications are presented in Table 2. The internal markers are black plastic beads that are colored differently compared to the tactile surface to ensure maximum color contrast for better visibility. The markers are arranged radially in the hemisphere at different uniform angles. Circular markers are non-directional and provide the same distinctive features in the image independent of motion direction [32]. To view the markers, a camera is mounted facing the backside of the tactile surface. When force is applied to the elastic surface of the sensor, the hemispherical surface deforms considerably, changing the magnitude and direction of the markers' displacement. To ensure good visibility, we illuminated the internal body enclosure with white LEDs. Our sensor fabrication process is similar to the making process in [37] with a difference in the markers' installation step, where their markers are painted protrusions on the internal surface of the elastic body.

3.3. VBTS contact pose dataset

Training and testing data for the tactile sensor was generated using a robotic manipulator and a calibrated flat surface of precisely known orientation and position as shown in Fig. 7-a. While some aircraft parts may have a slight curvature, this curvature is typically low, allowing most workpieces to be considered locally flat. The data collection routine with the measurements range is illustrated in Fig. 7-b, demonstrating the range of poses and orientations covered during data generation. The manipulator is commanded to orient the end-effector in spherical geometry by changing the roll θ and pitch ϕ angles. For each set of angles, the robotic manipulator proceeds to the contact surface and presses the tactile surface along its normal vector Z against the flat contact surface. The dataset includes the sensor orientation measured by the two normal angle components of the sensor tip (*i.e.*, roll θ and pitch ϕ angles) and a measure of how deep the tactile surface is pressed against a flat contact surface after achieving contact.

An image is sampled along with the normal angle and sensor depth at each variation of the roll or the pitch angles only after achieving contact with the flat surface. For each pose, five iterations are sampled, resulting in 2440 frames. To expand the dataset, frame augmentation was performed by rotating each frame three times by 90° each. Also, frame-shifting and scaling were performed to account for any slight changes in the sensor location relative to the vision sensor that may result from the excessive stress on the tactile surface. As pre-processing steps, the images are cropped to select the Region-of-Interest (ROI)

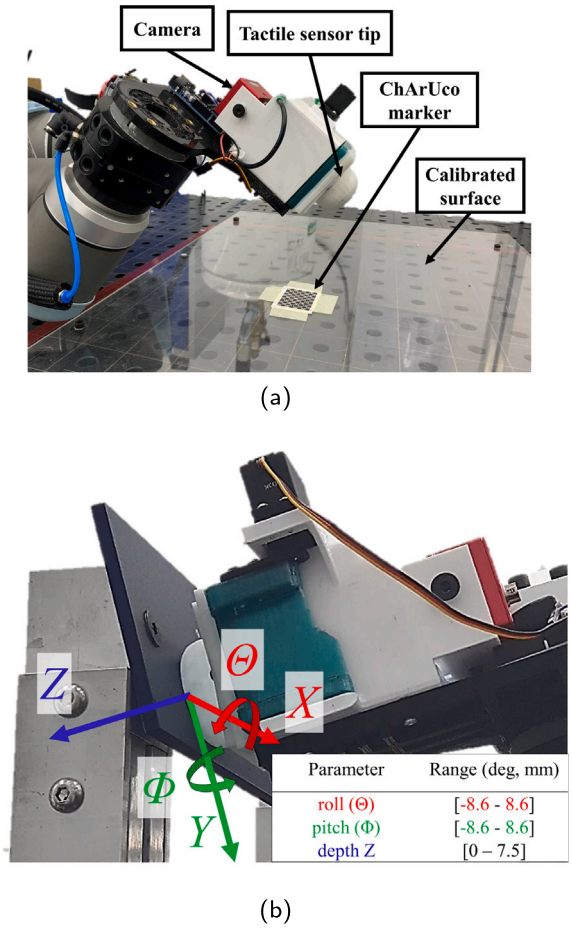

Fig. 7. The data collection setup with the tactile sensor mounted on a UR10 robot. (a) A calibrated flat surface is used for data collection, where the angles and depth are varied and the corresponding images are sampled from the camera feedback. (b) Roll θ, pitch ϕ and depth Z of the sensor frame relative to the workpiece surface are predicted. The table shows the measurements range for dataset collection used for training/testing the neural network.

and eliminate any pixels outside the inner surface of the sensor in the image. The ROI is resized to an image size of 100 × 100 pixels. Finally, the pixel intensity values are normalized to $[0, 1]$ using min–max normalization. The dataset was split in the ratio of 60:20:20, that is 60% data will go to the training set, 20% to the validation set and the remaining to the testing set.

3.4. Contact pose estimation using convolutional neural networks

Modern industries are being significantly transformed through the use of intelligent automation technologies. In particular, deep learning algorithms are rapidly getting involved in most of the fields of industry [53–55]. The use of deep learning is greatly accelerating the development of computer vision algorithms, which in turn directly contributes to the development of vision-based tactile sensor algorithms for various fields and applications [38,41,48]. In this subsection, we present our proposed Convolutional Neural Network to precisely estimate the relative contact pose between the machining tool and the workpiece in terms of normality and depth within the specifications of the aerospace industry. A CNN regression model with the architecture shown in Fig. 8 is trained using the dataset explained in Section 3.3. The input to the CNN is a 100 × 100 image. The CNN consists of N_{conv} convolutional layers, each layer contains two consecutive convolution operations with $N^i_{filters}$ filters and a kernel size of K^i_{size}, where $i \in [1, N_{conv}]$. Each convolutional layer is followed by a 2 × 2 max-pooling layer. The output of the last convolutional layer is then passed through a sequence of N_{dense} fully connected layers each having a width of N^j_{width}, where $j \in [1, N_{dense}]$. Each fully connected layer is followed by a ReLU activation function. The final layer contains three units corresponding to roll θ, pitch ϕ, and depth Z.

To select the best model hyperparameters, we run a Bayesian hyperparameter optimization process over the parameters shown in Table 3 following the guidelines and ranges explained in [38]. We limited the maximum range for the number of layers and filters to enable the CNN model to run in real-time on an embedded PC with limited computational power, which is critical for closed-loop control applications similar to ours. For each iteration of the Bayesian hyperparameter optimizer, the model weights are initialized using the built-in $\varepsilon glorot\_uniform\varepsilon$ weight initializer in Keras [56]. We selected the Mean Squared Error as the training loss function. The training and model optimization was implemented on a laptop with $RTX3080$ GPU running Windows 10. The resultant optimal CNN model parameters are also shown in Table 3.

3.5. Visual detection for robot guidance

This subsection presents the visual detection methods used for robot guidance and the 6-Dof localization of both the workpiece and its holes. Machine vision has been extensively employed in the literature for visual guidance to improve the end-effector's positional accuracy. Several studies, including [45,57], have explored 2D vision systems for enhancing positional accuracy. [12] combined 2D camera detection with laser sensors. In this study, we employ a dual-method approach: initially leveraging fiducials for global localization of the workpiece, followed by hole detection for precise refinement. In the first stage, ArUco fiducials are used to estimate the workpiece's pose. Using a priori estimate of the ${}_B\hat{T}_{Ar}$, the camera is placed facing the ArUco fiducial while the guidance/tactile sensing gate is opened; then it uses the ArUco detection method included in the OpenCV library [58] to refine ${}_B\hat{T}_{Ar}$. Since the transformation ${}_{Ar}T_{h^i}$ between each hole and the fiducial is known, an initial estimate of each hole's pose relative to the robot base can then be obtained as follows:

$$ {}_B\hat{T}_{h^i} = {}_B\hat{T}_{Ar}\,{}_{Ar}T_{h^i} \tag{3} $$

We utilize the camera's visual feedback to detect and localize circular holes to refine each hole's pose estimate to the required sub-millimeter accuracy. The robot first moves the camera to face each hole with a pre-defined standoff δZ using the priori ${}_B\hat{T}_{h^i}$. Once the camera is facing the hole, the robot refines the estimate ${}_B\hat{T}_{h^i}$ using the ED circle detector presented in [59]. This method is parameter-free, runs in real-time, and produces a low number of false positives. The ED circles detector outputs an array of all circular features in the image:

$$ c_k = \langle u_k, v_k, r_k \rangle, \qquad k \in [0, N_c] \tag{4} $$

where (u_k, v_k) are the pixel coordinates of the kth detected circle, r_k is the corresponding circle radius in pixel coordinates, and N_c is the overall number of detected circles. Since the camera's stand-off relative to the hole is known, these measurements can be converted to cartesian coordinates in meters using the pinhole model as follows:

$$ \begin{aligned} x_k &= \frac{f}{\delta Z}(u_k - c_x) \\ y_k &= \frac{f}{\delta Z}(v_k - c_y) \\ R_k &= \frac{f}{\delta Z} r_k \end{aligned} \tag{5} $$

where f is the camera's focal length, (c_x, c_y) are the camera's optical center, and (x_k, y_k) denote the cartesian coordinates of the kth circle.

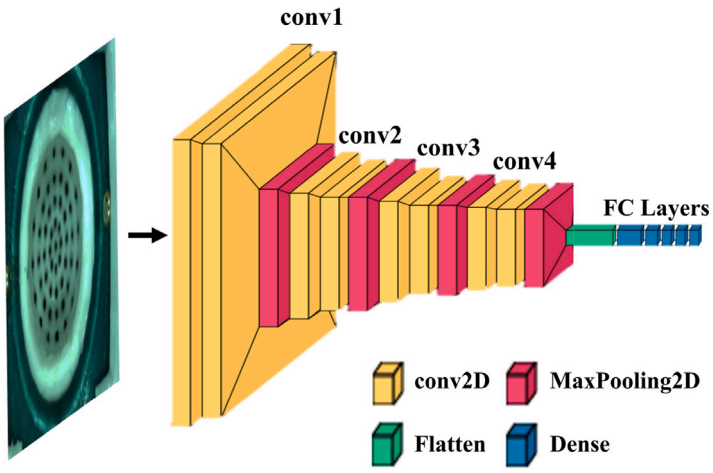

Fig. 8. The CNN architecture used in this work. A tactile image is processed by a sequence of convolutional, max pooling and fully connected layers.

Table 3
The ranges for the CNN hyperparameter optimization along with the corresponding optimal values.

Parameter	Parameter range	Optimal value
Number of convolutional layers N_{conv}	[1, 2, 3, 4]	4
Number of convolutional filters $N_{filters}$	[32, 64, 128, 256]	$\{N^1_{filters}, \dots N^4_{filters}\} = \{256, 256, 64, 32\}$
Size of convolutional kernel K_{size}	$[3\times3, 5\times5]$	$\{K^1_{size}, \dots K^4_{size}\} = \{5\times5, 5\times5, 3\times3, 3\times3\}$
Number of fully connected layers N_{dense}	[1, 2, 3, 4, 5]	4
Width of fully connected layers N_{width}	[16, 32, 64, 128, 256, 512]	$\{N^1_{width}, \dots N^4_{width}\} = \{256, 128, 64, 16\}$
Dropout coefficient	[0.0, 0.5]	0.0
Batch size	[64, 128, 256, 512]	512
Number of epochs	[128, 256, 512, 1024]	512

To filter out undesired circle detections, we calculate the probability of each circle being the target hole using prior knowledge of the circle's radius denoted by R^*, and assuming that the target hole should ideally be aligned with the camera's optical axis such that $x_k = y_k = 0$. We model any deviation from these nominal conditions as a multivariate normal distribution and calculate a score p_k for each circle detected as follows:

$$p_k = \frac{exp(-0.5(\chi_k - \mu)^T \Sigma^{-1}(\chi_k - \mu))}{\sqrt{(2\pi)^3|\Sigma|}} \tag{6}$$

where:

$$\chi_k = \begin{bmatrix} x_k \\ y_k \\ R_k \end{bmatrix} \tag{7}$$

and:

$$\mu = \begin{bmatrix} 0 \\ 0 \\ R^* \end{bmatrix} \tag{8}$$

and Σ is a tunable covariance matrix.

The circle with the highest score is then selected as:

$$k^* = \arg\max_k \quad p_k \tag{9}$$

and the hole's position vector relative to the camera is constructed as follows:

$$^C_C\vec{P}_{h^i} = \begin{bmatrix} x_{k^*} \\ y_{k^*} \\ \delta Z \end{bmatrix} \tag{10}$$

Finally, using the camera's pose $_BT_c$ obtained by solving the robot's forward kinematics, the hole's pose relative to the inertial base frame is calculated as:

$$_BT_{h^i} = {}_BT_C \ \ {}_CT_{h^i} \tag{11}$$

3.6. Robot guidance and control

Our proposed visual and tactile sensor was integrated into a full system with a cobot and a deburring end effector to experimentally verify the proposed design and algorithms. In this section, we will explain the control logic, including both position-based visual servoing (PBVS) for robot positioning and angle-based tactile servoing (ABTS) for normality adjustment. Sections 3.6.1 and 3.6.2 include comprehensive descriptions of both phases.

3.6.1. Position-based visual servoing (PBVS)

The PBVS stage directs the deburring end-effector towards initial alignment with the holes in the workpiece using the pose estimates obtained from visual feedback, as explained in Section 3.5. From the estimated hole pose $_BT_{h^i}$, we can define a desired end-effector pose $_B\bar{T}_E$ as follows:

$$_B\bar{T}_E = {}_B T_{h^i} \ \ {}_{h^i}\bar{T}_S \ \ ({}_ET_S)^{-1} \tag{12}$$

where $_{h^i}\bar{T}_S$ is the desired relative pose between the hole and the tool, and $_ET_S$ is the calibrated tool pose relative to the robot end-effector.

Accordingly, we define a desired joint angles vector $\hat{\theta} \in \mathbb{C}$ such that:

$$_B\bar{T}_E = g(\bar{\theta}) \tag{13}$$

Eq. (13) is solved to obtain $\bar{\theta}$ using the Newton–Raphson inverse kinematic approach of the open-source Kinematic and Dynamics Library (KDL).[4] We then generate a joint trajectory $\bar{\theta}(t)$ between the current joint angles and the desired ones using the implementation of RRT-connect [60] of the Open Motion Planning Library [61]. A low-level PID controller then controls each joint to follow $\bar{\theta}(t)$.

A control loop diagram for position-based robot control is shown in Fig. 9. This process explained above is repeated multiple times until

[4] KDL: https://www.orocos.org/kdl.html.

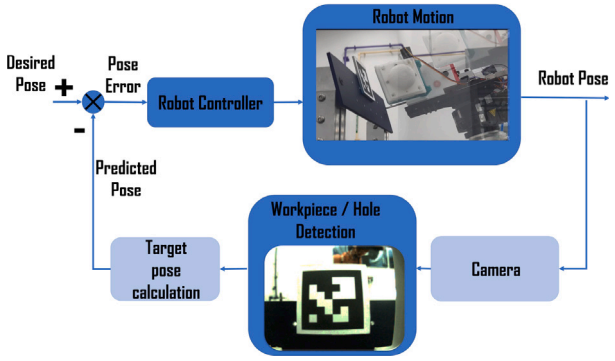

Fig. 9. Position-based robot control loop.

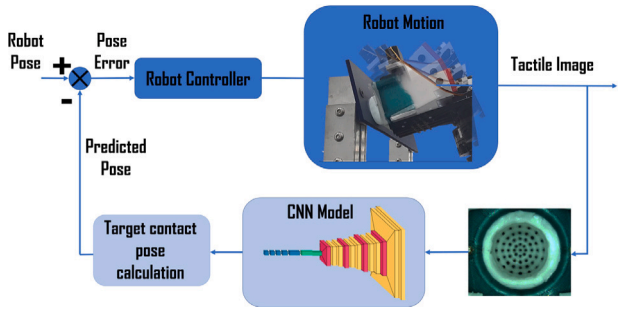

Fig. 10. The angle-based tactile servoing control loop.

the pose error converges. The pose error is defined as the difference between the current end-effector pose and the desired one, and is denoted by: $E = ({}_B T_E)^{-1} \ {}_B\bar{T}_E$.

3.6.2. Angle-based tactile control

This stage of the controller is concerned with the fine refinement of the deburring tool's pose to guarantee perpendicularity and precise depth during the machining process. It uses the contact pose estimates from the CNN to refine the alignment of the tool. Once the tactile surface achieves contact with the workpiece, the CNN outputs roll θ, pitch ϕ, and depth Z estimates. We define the rotation vector $\vec{r} = [\theta, \phi, 0]^T$ between the end-effector and the workpiece. This rotation vector can be converted to a 3×3 rotation matrix using the Rodrigues rotation formula as follows:

$$ {}_T R_W = \begin{bmatrix} \theta^2\sigma + cos(\|\vec{r}\|) & \theta\phi\sigma & -\phi sin(\|\vec{r}\|) \\ \phi\theta\sigma & \phi^2\sigma + cos(\|\vec{r}\|) & \theta sin(\|\vec{r}\|) \\ \phi sin(\|\vec{r}\|) & -\theta sin(\|\vec{r}\|) & cos(\|\vec{r}\|) \end{bmatrix} \tag{14} $$

where $\sigma = 1 - cos(\|\vec{r}\|)$.

Additionally, depth measurements Z can be used to refine the priori estimate of the workpiece's position relative to the tactile sensor denoted by ${}^T_T\hat{\bar{P}}_W$, as follows:

$$ {}^T_T\hat{P}^+_W = {}^T_T\hat{\bar{P}}_W + \begin{bmatrix} 0 \\ 0 \\ Z \end{bmatrix} \tag{15} $$

which can be used alongside Eq. (14) to construct the transformation matrix ${}_T T_W$. The pose of the workpiece relative to the inertial base frame can hence be defined as:

$$ {}_B T_W = {}_B T_E \ {}_E T_T \ {}_T T_W \tag{16} $$

Similarly, the hole's pose is defined as:

$$ {}_B T_{h^i} = {}_B T_W \ {}_W T_{h^i} \tag{17} $$

where ${}_W T_{h^i}$ is known from a CAD model of the workpiece.

Table 4
Contact pose prediction performance of the proposed CNN on an unseen testing subset.

Parameter	MAE	Standard deviation
Angle Error (deg)	0.13	0.14
depth Z (mm)	0.3	0.3

Prior to initiating the deburring process, the deburring tool needs to be at a specified normal pose relative to the hole. We denote this desired pose by ${}_{h^i}\bar{T}_S$, which can be used to define a desired end-effector pose as:

$$ {}_B\bar{T}_E = {}_B T_{h^i} \ {}_{h^i}\bar{T}_S \ ({}_E T_S)^{-1} \tag{18} $$

This end-effector pose is then achieved in a similar manner to the PBVS case explained in Section 3.6.1. The control loop of the angle-based control is shown in Fig. 10.

4. Results and discussion

The presented multi-functional vision-based sensor and the related algorithms were tested on the robotic setup shown in Section 2. Our experimental evaluation focuses on two main aspects. At first, we assess the contact pose estimation using the proposed CNN in terms of Mean Absolute Error (MAE) for both the angle and depth estimates, calculated as follows:

$$ MAE = \frac{\sum_{i=1}^{n} |X^i_{prediction} - X^i_{actual}|}{n} \tag{19} $$

where X can be the roll θ, pitch ϕ, the square sum of roll θ and pitch ϕ or the depth Z. And n is the number of samples in the testing subset. It is important to note that the ground truth data (X_{actual}) used for evaluation was taken from the same dataset which involved the calibrated flat surface for data collection as explained in Section 3.3.

Next, we present the evaluation data for the full performance of the system in a deburring application. The sequence of operations to perform deburring tasks utilizing our multi-functional vision-based sensor consists of several phases. Starting with the exploration phase and the workpiece and hole detection phase to localize the workpiece in 6 DoF and align the end-effector with the hole. Then the orientation and depth detection and correction phase, where the cobot achieves contact with the workpiece and uses the proposed CNN algorithm to estimate the contact properties in terms of orientation and depth and correct for them. The last stage is the deburring phase, in which we perform deburring for both sides of the workpiece. The overall sequence of operation is depicted in Fig. 11.

The performance of the developed CNN for estimating the normality and contact pose was tested on an unseen subset of the collected data. In Fig. 12, we present a sample of testing images with the corresponding ground truth and CNN predictions. Table 4 summarizes the obtained results in terms of MAE and standard deviation of the contact pose, across all the images in the testing subset acquired as described in Section 3.3. Table 4 shows that our proposed vision-based tactile sensor is capable of estimating the normal axis of the workpiece with an MEA of 0.13° . Furthermore, the sensor can estimate the relative depth between the end-effector and the workpiece with an MAE of 0.3 mm. The maximum normality estimation error achieved on the testing set was 1.18°, which conforms with the perpendicularity requirements of the aerospace manufacturing industry, and is well below the specified error tolerance of 2° . The developed perception algorithms are also suitable for real-time operation without the need for a GPU, with the inference time running on an intel i5-10210U processor being 0.055 s, which corresponds to a frame rate of approximately 18 fps.

In Table 5, we benchmark our results against state-of-the-art in vision-based tactile sensing. The results clearly show that our sensor and perception algorithms outperform prior studies in the literature and achieve over 50% reduction in error compared to these studies.

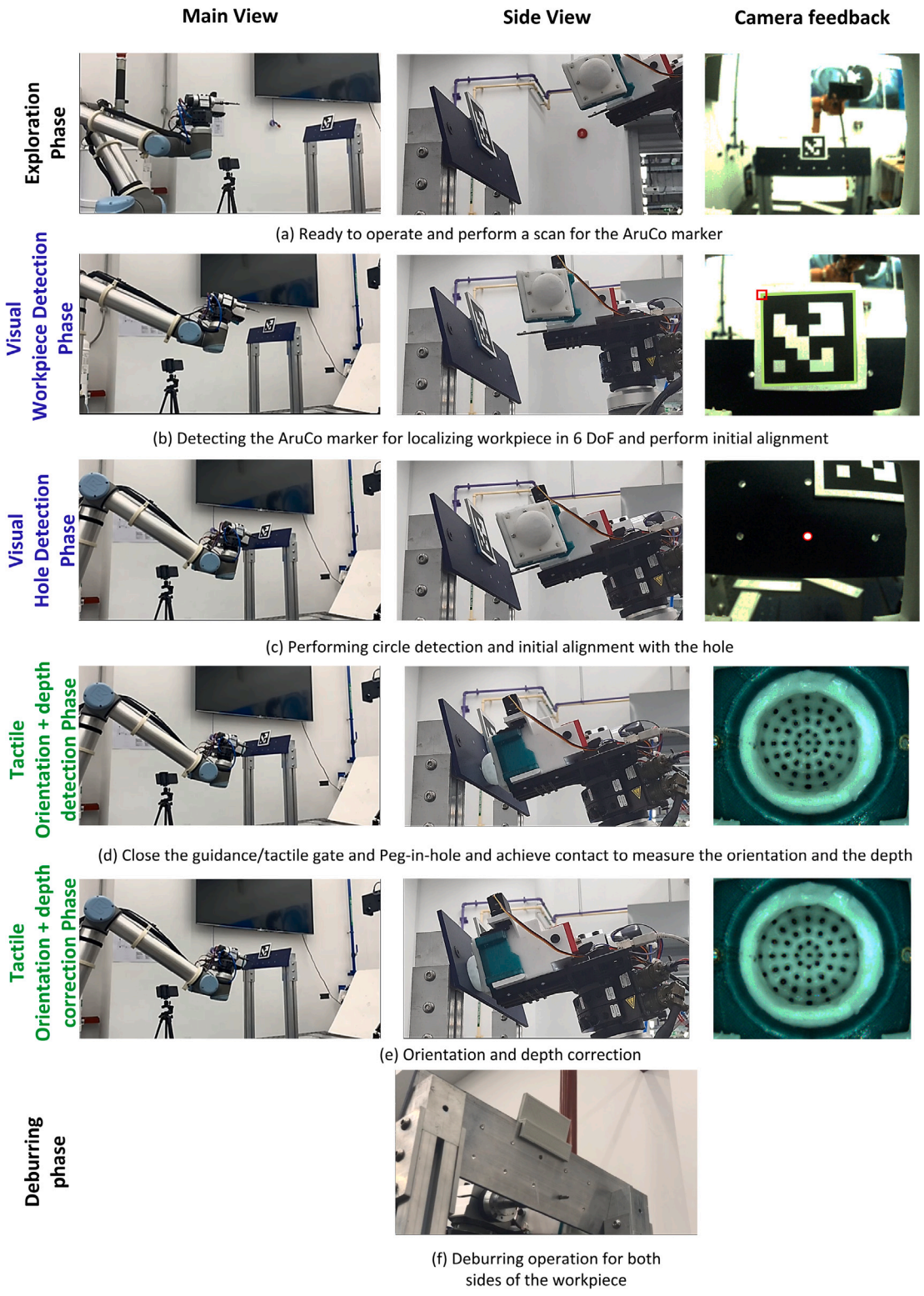

Fig. 11. The sequence of operation to perform deburring tasks utilizing our multi-functional vision-based sensor. In the exploration phase and the workpiece detection phase, the sensor enables the cobot to localize the workpiece in 6-DoF. In the hole detection phase, the cobot performs circle detection for initial alignment with the hole. In the orientation and depth detection phase, the guidance/tactile sensing gate is closed, and the robot achieves contact with the workpiece and uses CNN to estimate the contact properties in terms of normality and depth. Then the orientation and depth are corrected, and the deburring phase starts.

In addition to quantitative advantages, our sensor is the only one that provides exteroceptive perception in addition to proprioceptive tactile sensing, while the capabilities of all prior work is limited to tactile sensing only.

We have tested and validated our sensor in the automated deburring setup shown in Fig. 1. The cobot deburring process is performed in two main stages: the visual workpiece/hole localization stage and the tactile orientation/depth measurement and correction stage. At first,

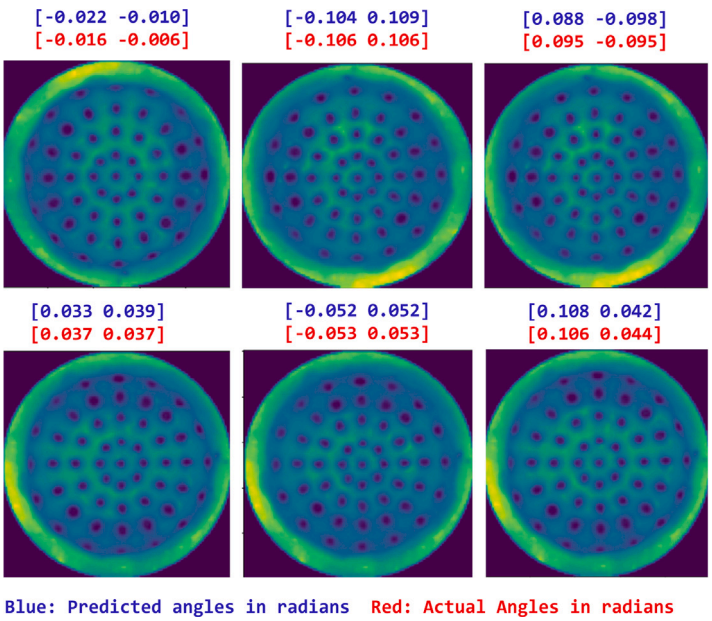

Fig. 12. The camera feedback of the tactile sensor at different random contact poses with the workpiece depicted from the testing subset. The actual benchmark roll and pitch angle values (in radians) are shown with the predicted values from the CNN model.

Table 5
Comparison between our multi-functional vision-based sensor and other works in the literature in terms of contact pose estimation. The results show that our sensor and the corresponding perception algorithms outperform prior work in the literature in terms of MAE of contact angles.

Studies	MAE of contact angles	Sensor functionality	Comments
A. Gupta et al. [44]	Yaw: 2.13°	limited to proprioceptive perception	The sensor only measures yaw angle
E. Pso. et al. [42]	Roll: 0.33° Pitch: 0.28°	limited to proprioceptive perception	The sensor is used for object grasping task
N. Lepora et al. [38]	Roll: 0.3° Pitch: 0.3°	limited to proprioceptive perception	The sensor is used for 3D surfaces pose contact estimation with shear motion induced.
This work	**Roll: 0.13° Pitch: 0.13°**	**exteroceptive and proprioceptive perception**	A multi-functional sensor for exteroceptive and proprioceptive perception that meets the aircraft industry requirements

our multi-functional sensor gate is opened, and the cobot performs 6-DoF localization of the workpiece and its holes using the ArUco fiducial. Afterwards the cobot uses circular hole detection and PBVS for initial positional alignment with each hole. The sensor gate is then closed as the deburring tool is inserted in the hole. The cobot utilizes VBTS and our trained CNN model to estimate the contact pose, and refines the normality and depth of contact using angle-based tactile control. Fig. 13 illustrates the normality and depth refinement stages of the deburring process. Once the tool is properly aligned with the hole, the deburring motor is finally activated to deburr both sides of the workpiece. Fig. 14 shows the holes before and after they were deburred using our system. The control schemes of the deburring experiments are shown in Fig. 15.

The deburring experiments were conducted on two workpieces placed differently in the environment with ten holes in each. We evaluate the repeatability of the overall system in terms of the counter-sink diameter resulting from the deburring of the holes, measured using Olympus BX51 microscope, as explained in Fig. 16. Measuring the counter-sink diameter provides a simple and easy-to-quantify metric, especially since the diameter and depth of a countersink are directly correlated when the countersink angle is known. This method avoids the need for specialized gauges or laser profilometry, making it a viable option for assessing system performance. Table 6 presents the counter-sink diameter for all the holes on both workpieces. The results show a standard deviation (std) for the measured counter sink diameters of less than 0.2 mm, which satisfies the precision requirements of most deburring processes in the aerospace industry. In particular, the evaluation of deburred holes in the industry is usually conducted based on the difference between the minimum and maximum countersink diameters. In our experiments, we observed the minimum and maximum values to be 5.864 and 6.489 respectively, with a difference of 0.485 mm. While there is room for improving the precision in terms of countersink diameter, these errors can arise from multiple sources that are not within the scope of this paper and are not necessarily measurement errors. These errors include vibrations in the deburring motor, tool misalignment with the motor's center of rotation, and the lack of clamping force to stabilize the end-effector. These sources of error can be addressed by a better end-effector design that stabilizes the end-effector and minimizes the effects of the low stiffness in the cobot's joints.

In terms of assessing the positioning accuracy of our system, we note that the radius of the holes on the workpiece is $2.55\ \text{mm}^{+0.05}_{-0.05}$, while the deburring tool has a radius of 2.35 mm. As such, the maximum allowable positioning error during the deburring pin insertion process is 0.2 mm, and any errors beyond that would cause a failure in insertion. The success rate of the insertion process in our experiments was 100%, which shows that the maximum positioning errors of our system do not exceed 0.2 mm, which meets the positional accuracy requirement of the deburring process in the aerospace industry of less than 0.5 mm [11,12]. State-of-the-art studies that utilize exteroceptive perception technologies to improve the machining position accuracy of a robotic system achieved positioning accuracy ranging from 0.05 mm to 0.3 mm [12,45,57,62]. Nonetheless, our work is the only one to utilize a cobot for machining instead of an industrial robot. Both the normality error and the positioning error conform with the aerospace standards with an MAE of 0.13° and an accuracy of less than 0.2 mm, respectively.

Table 6
Counter sink diameter in mm for the holes' deburring experiments across multiple workpieces. The counter sink diameter standard deviation complies with the insertion precision needed of less than 0.2 mm.

Part	Hole ID (mm)										Mean	Std	Min	Max
	1	2	3	4	5	6	7	8	9	10				
A	6.215	6.121	5.968	6.413	6.315	6.223	6.015	6.416	5.864	6.489	6.204	0.199	5.864	6.489
B	6.426	6.374	6.264	6.219	6.264	6.149	6.164	5.984	6.058	6.079	6.1981	0.132	5.984	6.426
										Aggregate	**6.201**	**0.166**	**5.864**	**6.489**

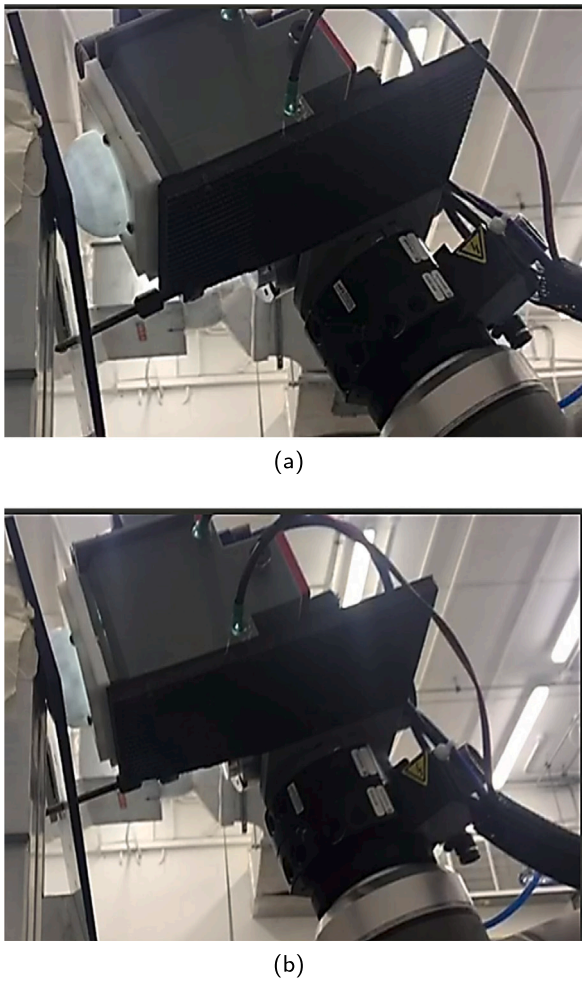

Fig. 13. (a) The end-effector before the normal alignment while the sensor is pressing on the workpiece. (b) The end-effector after the normal alignment.

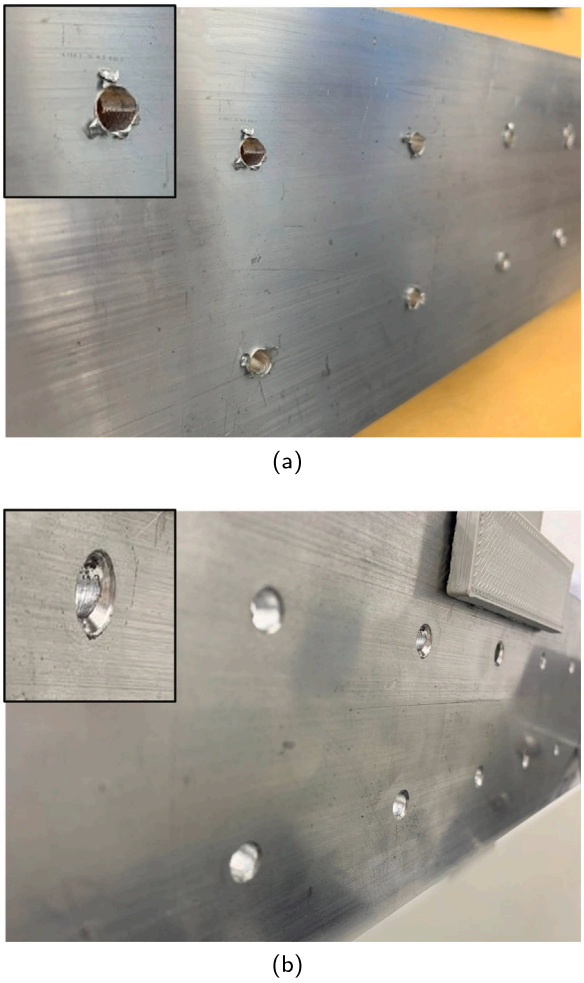

Fig. 14. (a) The backside of the target workpiece **before the deburring process**. (b) The backside of the target workpiece **after the deburring process** using the proposed vision-based tactile sensor for normal alignment.

Our multi-functional vision-based sensor and the corresponding perception algorithms have the potential to be used in a wide range of industrial applications. For instance, the sensor can be used in robotic drilling applications, where precise alignment and normal orientation are crucial to ensure accurate drilling and minimal material damage. Our sensor addresses both of these requirements, providing high positioning and normality measurement accuracy, which greatly enhances the drilling quality to meet aerospace standards. The integration process for these applications can be done seamlessly, with minimal changes to the current configuration by simply placing the drilling tool beside our proposed sensor. This allows for efficient and precise robotic drilling operation, leading to improved production quality and efficiency. Furthermore, the sensor can also be used in robotic riveting applications, where the same high positioning accuracy and good normality are needed to ensure a strong and reliable joint. Additionally, the sensor can also be used in robotic grinding and polishing applications, where precise contact properties and real-time control are required to achieve smooth and uniform surface finishes. The sensor can also be used in robotic assembly applications, where it can be used to ensure precise alignment and contact between different components. These are just a few examples of the various applications that our multi-functional vision-based sensor can be used in. The sensor's ability to provide real-time feedback on contact properties, as well as its ability to use both visual and tactile sensing, make it a versatile and powerful tool for a wide range of industrial applications.

5. Conclusions

This paper presented the first multi-functional vision-based sensor for precise positioning and normality control in robotic manufacturing. Our sensor introduces a unique design that enables both exteroceptive

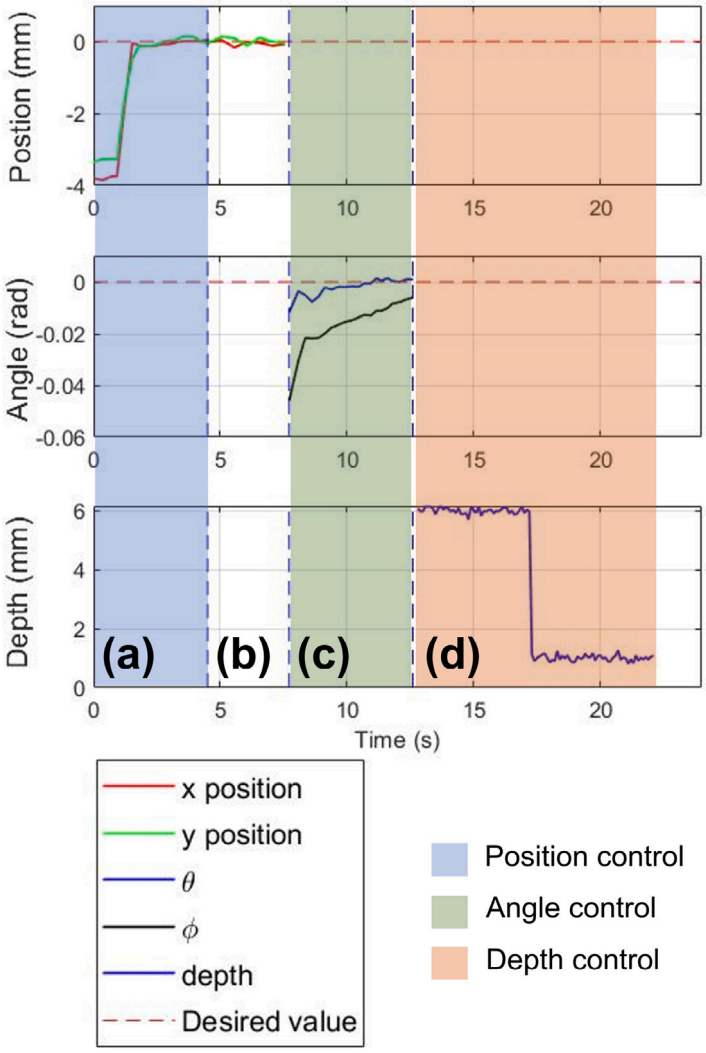

Fig. 15. The control stages of the deburring experiments. (a) After performing initial alignment with the workpiece and while our multi-functional sensor gate is opened, the robotic deburring experiment starts by performing hole detection to precisely localize the hole by position-based visual servoing. (b) In this stage, the multi-functional sensor gate is closed and Peg-in-hole is performed. (c) In this stage the orientation is precisely corrected and the deburring tool pose is controlled using the angle-based tactile control scheme. (d) The last stage guarantees precise depth control for precise deburring of both sides of the workpiece.

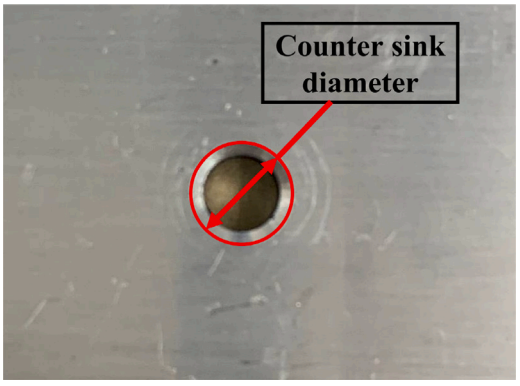

Fig. 16. The counter sink diameter measurement used to assess the precision of the deburring process.

and proprioceptive perception using a single imaging sensor. Exteroceptive perception is used to position the robotic end-effector relative to an arbitrarily placed workpiece with sub-millimeter precision. As for proprioceptive perception, our sensor utilizes a low-cost deformable surface and Convolutional Neural Networks to estimate and adjust the contact pose between the end-effector and the workpiece. We also introduced the visual and tactile control laws that facilitate the utilization of our sensor in robotic manufacturing scenarios.

We have validated our sensor experimentally for a deburring application using a cobot. The use of a cobot enables the system to be deployed directly in the production line alongside human technicians working on different machining tasks. Our tests verify that the proposed vision-based sensor provides the high precision required in the aerospace manufacturing industry for both positioning and normality control. Quantitative results show an average normality error of 0.13°, and positioning errors below 0.2 mm, both of which are within the allowed error tolerances in the aerospace industry. Therefore, our multi-functional sensor establishes a new paradigm for perception systems in robotic manufacturing, as it enhances current methods in terms of precision, cost, weight, size, and low instrumentation and wiring requirements. The ability to provide both positioning and normality measurements using a single modality has the potential to accelerate the development of autonomous manufacturing systems, as opposed to traditional practices of developing an independent sensing sub-system for each of these functionalities. Furthermore, the extension of VBTS

capabilities to perceive external environments in addition to providing tactile measurements would lead to higher adoption of VBTS within the robotics community, especially since most robotic applications require a combination of visual and tactile perception.

While our analysis and experiments in this paper validate the sensor for real-world applications and prove that the novel sensor meets the specifications of the aerospace industry, there is room for improvement and research to enhance the sensor and increase its range of functions. One possible improvement is to enhance the optical system to augment its robustness to variations in lighting conditions. For instance, when the robot is performing exteroceptive perception in low-light conditions, conventional frame-based cameras suffer from motion blur and high latency due to increased exposure timing. On the other hand, when the guidance/tactile sensing gate is closed for proprioceptive perception, minimal external light enters the sensor, which adds the requirement of an internal lighting source. Other areas of improvement include extending the lifetime of the elastic tactile interface to prevent plastic deformation over long-term operation. We will address those limitations in future work, which would include the following:

- Enhance the robustness of lighting conditions and eliminate the need for an internal lighting source by using neuromorphic vision.
- Extend the functions of the vision-based tactile sensor by adding vibration and force/torque estimation capabilities, which might require higher frame rate camera or neuromorphic camera.
- Optimize the tactile sensor material and design for enhanced measurement range, sensitivity, and sensor lifetime and explore miniaturization to assess the impact on accuracy.
- Optimize the sensor design in terms of weight and size.

CRediT authorship contribution statement

Mohamad Halwani: Methodology, Software, Investigation, Design, Writing. **Abdulla Ayyad:** Conceptualization, Methodology, Software, Experimentation, Writing. **Laith AbuAssi:** Conceptualization, End-effector design, Software. **Yusra Abdulrahman:** Methodology, Technical advising, Review and editing. **Fahad Almaskari:** Supervision, Review and editing. **Hany Hassanin:** Project management, Review and editing. **Abdulqader Abusafieh:** Project management, Review and editing. **Yahya Zweiri:** Project management, Funding acquisition, Review and editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Abdulla Ayyad has patent pending to Aerospace Holding Co LLC Khalifa University of Science, Technology and Research (KUSTAR).

References

[1] Welber I. Factory of the future. IEEE Control Syst Mag 1987;7(2):20–2. http://dx.doi.org/10.1109/MCS.1987.1105295.

[2] Huang X, Halwani M, Muthusamy R, Ayyad A, Swart D, Seneviratne L, et al. Real-time grasping strategies using event camera. J Intell Manuf 2022;33(2):593–615. http://dx.doi.org/10.1007/s10845-021-01887-9.

[3] Perez-Vidal C, Gracia L, de Paco JM, Wirkus M, Azorin JM, de Gea J. Automation of product packaging for industrial applications. Int J Comput Integr Manuf 2018;31(2):129–37. http://dx.doi.org/10.1080/0951192X.2017.1369165.

[4] Oztemel E, Gursev S. Literature review of industry 4.0 and related technologies. J Intell Manuf 2020;31(1):127–82. http://dx.doi.org/10.1007/s10845-018-1433-8.

[5] Verl A, Valente A, Melkote S, Brecher C, Ozturk E, Tunc LT. Robots in machining. CIRP Ann 2019;68(2):799–822. http://dx.doi.org/10.1016/j.cirp.2019.05.009, URL https://www.sciencedirect.com/science/article/pii/S0007850619301660.

[6] Eguti CCA, Trabasso LG. Design of a robotic orbital driller for assembling aircraft structures. Mechatronics 2014;24(5):533–45. http://dx.doi.org/10.1016/j.mechatronics.2014.06.007, URL https://www.sciencedirect.com/science/article/pii/S0957415814000968.

[7] Karim A, Verl A. Challenges and obstacles in robot-machining. In: IEEE ISR 2013. 2013, http://dx.doi.org/10.1109/ISR.2013.6695731.

[8] Chen Y, Dong F. Robot machining: Recent development and future research issues. Int J Adv Manuf Technol 2013;66(9):1489–97. http://dx.doi.org/10.1007/s00170-012-4433-4.

[9] Chen D, Yuan P, Wang T, Cai Y, Tang H. A normal sensor calibration method based on an extended Kalman filter for robotic drilling. Sensors (Switzerland) 2018;18(10). http://dx.doi.org/10.3390/s18103485.

[10] Yu L, Bi Q, Ji Y, Fan Y, Huang N, Wang Y. Vision based in-process inspection for countersink in automated drilling and riveting. Precis Eng 2019;58:35–46. http://dx.doi.org/10.1016/j.precisioneng.2019.05.002, URL https://www.sciencedirect.com/science/article/pii/S0141635919300613.

[11] Yuan P, Wang Q, Wang T, Wang C, Song B. Surface normal measurement in the end effector of a drilling robot for aviation. In: Proceedings - IEEE international conference on robotics and automation. IEEE; 2014, p. 4481–6. http://dx.doi.org/10.1109/ICRA.2014.6907513.

[12] Frommknecht A, Kuehnle J, Effenberger I, Pidan S. Multi-sensor measurement system for robotic drilling. Robot Comput-Integr Manuf 2017;47. http://dx.doi.org/10.1016/j.rcim.2017.01.002.

[13] Santos KRDS, De Carvalho GM, Tricarico RT, Ferreira LFL, Villani E, Suterio R. Evaluation of perpendicularity methods for a robotic end effector from aircraft industry. In: 2018 13th IEEE international conference on industry applications, INDUSCON 2018 - Proceedings. 2019, p. 1373–80. http://dx.doi.org/10.1109/INDUSCON.2018.8627218.

[14] Yu L, Zhang Y, Bi Q, Wang Y. Research on surface normal measurement and adjustment in aircraft assembly. Precis Eng 2017;50:482–93. http://dx.doi.org/10.1016/j.precisioneng.2017.07.004.

[15] Zhang Y, Bi Q, Yu L, Wang Y. Online adaptive measurement and adjustment for flexible part during high precision drilling process. Int J Adv Manuf Technol 2017;89(9):3579–99. http://dx.doi.org/10.1007/s00170-016-9274-0.

[16] Lin M, Yuan P, Tan H, Liu Y, Zhu Q, Li Y. Improvements of robot positioning accuracy and drilling perpendicularity for autonomous drilling robot system. In: 2015 IEEE international conference on robotics and biomimetics, IEEE-ROBIO 2015. IEEE; 2015, p. 1483–8. http://dx.doi.org/10.1109/ROBIO.2015.7418980.

[17] Yuan P, Lai T, Li Y, Han W, Lin M, Zhu Q, et al. The attitude adjustment algorithm in drilling end-effector for aviation. Adv Mech Eng 2016;8(1):1–9. http://dx.doi.org/10.1177/1687814016629348.

[18] Shaomin L, Deyuan Z, yanqiang L, chunjian L, Hui T, guang M. A self-adaption normal direction and active variable stiffness low-frequency vibration-assisted system for curved surface drilling. Precis Eng 2020;64:307–18. http://dx.doi.org/10.1016/j.precisioneng.2020.04.017, URL https://www.sciencedirect.com/science/article/pii/S0141635920302233.

[19] Xiao R, Xu Y, Hou Z, Chen C, Chen S. An automatic calibration algorithm for laser vision sensor in robotic autonomous welding system. J Intell Manuf 2022;33(5):1419–32. http://dx.doi.org/10.1007/s10845-020-01726-3.

[20] Ayyad A, Halwani M, Swart D, Muthusamy R, Almaskari F, Zweiri Y. Neuromorphic vision based control for the precise positioning of robotic drilling systems. Robot Comput-Integr Manuf 2023;79:102419. http://dx.doi.org/10.1016/j.rcim.2022.102419, URL https://www.sciencedirect.com/science/article/pii/S0736584522001041.

[21] Furtado LFF, Coracini GK, Villani E, Trabasso LG. Comparative Study Between Two Methods for Perpendicularity Corrections in Robotic Manipulators. In: 21st international congress of mechanical engineering - COBEM, vol. 5, 2011, p. 1194–200.

[22] Tian W, Zhou W, Zhau W, Liao W, Zeng Y. Auto-normalization algorithm for robotic precision drilling system in aircraft component assembly. Chin J Aeronaut 2013;26(2):495–500. http://dx.doi.org/10.1016/j.cja.2013.02.029.

[23] Gao Y, Wu D, Nan C, Chen K. Normal direction measurement in robotic drilling and precision calculation. Int J Adv Manuf Technol 2015;76(5–8):1311–8. http://dx.doi.org/10.1007/s00170-014-6320-7.

[24] Gao Y, Wu D, Dong Y, Ma X, Chen K. The method of aiming towards the normal direction for robotic drilling. Int J Precis Eng Manuf 2017;18(6):787–94. http://dx.doi.org/10.1007/s12541-017-0094-4.

[25] Yu L, Zhang Y, Bi Q, Wang Y. Research on surface normal measurement and adjustment in aircraft assembly. Precis Eng 2017;50:482–93. http://dx.doi.org/10.1016/j.precisioneng.2017.07.004, URL https://www.sciencedirect.com/science/article/pii/S0141635917300491.

[26] Song T, Xi F, Guo S, Ming Z, Lin Y. A comparison study of algorithms for surface normal determination based on point cloud data. Precis Eng 2015;39:47–55. http://dx.doi.org/10.1016/j.precisioneng.2014.07.005, URL https://www.sciencedirect.com/science/article/pii/S014163591400124X.

[27] Rao G, Yang X, Yu H, Chen K, Xu J. Fringe-projection-based normal direction measurement and adjustment for robotic drilling. IEEE Trans Ind Electron 2020;67(11):9560–70. http://dx.doi.org/10.1109/TIE.2019.2952791.

[28] Olsson T, Haage M, Kihlman H, Johansson R, Nilsson K, Robertsson A, et al. Cost-efficient drilling using industrial robots with high-bandwidth force feedback. Robot Comput-Integr Manuf 2010;26(1):24–38. http://dx.doi.org/10.1016/j.rcim.2009.01.002.

[29] Shah UH, Muthusamy R, Gan D, Zweiri Y, Seneviratne L. On the design and development of vision-based tactile sensors. J Intell Robot Syst 2021;102(4):82. http://dx.doi.org/10.1007/s10846-021-01431-0.

[30] Lambeta M, Chou PW, Tian S, Yang B, Maloon B, Most VR, et al. DIGIT: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation. IEEE Robot Autom Lett 2020;5(3):3838–45. http://dx.doi.org/10.1109/LRA.2020.2977257, arXiv:2005.14679.

[31] hyun Choi S, Tahara K. Dexterous object manipulation by a multi-fingered robotic hand with visual-tactile fingertip sensors. ROBOMECH J 2020;7(1). http://dx.doi.org/10.1186/s40648-020-00162-5.

[32] Kumagai K, Shimonomura K. Event-based tactile image sensor for detecting spatio-temporal fast phenomena in contacts. In: 2019 IEEE world haptics conference. Institute of Electrical and Electronics Engineers Inc.; 2019, p. 343–8. http://dx.doi.org/10.1109/WHC.2019.8816132.

[33] Product manual syntouch biotac® SP tactile sensor. Tech. rep., 2020.

[34] Ito Y, Kim Y, Obinata G. Robust slippage degree estimation based on reference update of vision-based tactile sensor. IEEE Sens J 2011;11(9):2037–47. http://dx.doi.org/10.1109/JSEN.2010.2104316.

[35] Rigi A, Baghaei Naeini F, Makris D, Zweiri Y. A novel event-based incipient slip detection using dynamic active-pixel vision sensor (DAVIS). Sensors 2018;18(2). http://dx.doi.org/10.3390/s18020333, URL https://www.mdpi.com/1424-8220/18/2/333.

[36] Sun H, Kuchenbecker KJ, Martius G. A soft thumb-sized vision-based sensor with accurate all-round force perception. Nat Mach Intell 2022;4(2):135–45. http://dx.doi.org/10.1038/s42256-021-00439-3.

[37] Kakani V, Cui X, Ma M, Kim H. Vision-based tactile sensor mechanism for the estimation of contact position and force distribution using deep learning. Sensors 2021;21(5). http://dx.doi.org/10.3390/s21051920, URL https://www.mdpi.com/1424-8220/21/5/1920.

[38] Lepora NF, Lloyd J. Optimal deep learning for robot touch: Training accurate pose models of 3D surfaces and edges. IEEE Robot Autom Mag 2020;27(2):66–77. http://dx.doi.org/10.1109/MRA.2020.2979658.

[39] Sajwani H, Ayyad A, Alkendi Y, Halwani M, Abdulrahman Y, Abusafieh A, et al. TactiGraph: An asynchronous graph neural network for contact angle prediction using neuromorphic vision-based tactile sensing. Sensors 2023;23(14). http://dx.doi.org/10.3390/s23146451, URL https://www.mdpi.com/1424-8220/23/14/6451.

[40] Ward-Cherrier B, Pestell N, Lepora NF. NeuroTac: A neuromorphic optical tactile sensor applied to texture recognition. In: Proceedings - IEEE international conference on robotics and automation, 2020, p. 2654–60. http://dx.doi.org/10.1109/ICRA40945.2020.9197046, arXiv:2003.00467.

[41] Kakani V, Cui X, Ma M, Kim H. Vision-based tactile sensor mechanism for the estimation of contact position and force distribution using deep learning. Sensors 2021;21(5). http://dx.doi.org/10.3390/s21051920, URL https://www.mdpi.com/1424-8220/21/5/1920.

[42] Psomopoulou E, Pestell N, Papadopoulos F, Lloyd J, Doulgeri Z, Lepora NF. A robust controller for stable 3D pinching using tactile sensing. IEEE Robot Autom Lett 2021;6(4):8150–7. http://dx.doi.org/10.1109/LRA.2021.3104057.

[43] Muthusamy R, Huang X, Zweiri Y, Seneviratne L, Gan D, Muthusamy R. Neuromorphic event-based slip detection and suppression in robotic grasping and manipulation. IEEE Access 2020;8:153364–84. http://dx.doi.org/10.1109/ACCESS.2020.3017738, arXiv:2004.07386.

[44] Gupta AK, Aitchison L, Lepora NF. Tactile image-to-image disentanglement of contact geometry from motion-induced shear. In: Proceedings of the 5th conference on robot learning. Proceedings of machine learning research, vol. 164, PMLR; 2022, p. 14–23, URL https://proceedings.mlr.press/v164/gupta22a.html.

[45] Mei B, Zhu W. Accurate positioning of a drilling and riveting cell for aircraft assembly. Robot Comput-Integr Manuf 2021;69:102112. http://dx.doi.org/10.1016/j.rcim.2020.102112, URL https://www.sciencedirect.com/science/article/pii/S0736584520303227.

[46] Jiang T, Cui H, Cheng X, Tian W. A measurement method for robot peg-in-hole prealignment based on combined two-level visual sensors. IEEE Trans Instrum Meas 2021;70:1–12. http://dx.doi.org/10.1109/TIM.2020.3026802.

[47] AYYAD A. Robotic manipulator with visual guidance and tactile sensing (US 2023/0073681 A1. United States Patent and Trademark Office, March 2023).

[48] Macdonald FLA, Lepora NF, Conradt J, Ward-Cherrier B. Neuromorphic tactile edge orientation classification in an unsupervised spiking neural network. Sensors 2022;22(18). http://dx.doi.org/10.3390/s22186998, URL https://www.mdpi.com/1424-8220/22/18/6998.

[49] Dornaika F, Horaud R. Simultaneous robot-world and hand-eye calibration. IEEE Trans Robot Autom 1998;14(4):617–22. http://dx.doi.org/10.1109/70.704233.

[50] Ward-Cherrier B, Cramphorn L, Lepora NF. Exploiting sensor symmetry for generalized tactile perception in biomimetic touch. IEEE Robot Autom Lett 2017;2(2):1218–25. http://dx.doi.org/10.1109/LRA.2017.2665692.

[51] Sferrazza C, D'Andrea R. Design, motivation and evaluation of a full-resolution optical tactile sensor. Sensors (Switzerland) 2019;19(4). http://dx.doi.org/10.3390/s19040928.

[52] Ward-Cherrier B, Cramphorn L, Lepora NF. Tactile manipulation with a TacThumb integrated on the open-hand M2 gripper. IEEE Robot Autom Lett 2016;1(1):169–75. http://dx.doi.org/10.1109/LRA.2016.2514420.

[53] Liu C, Wang K, Wang Y, Yuan X. Learning deep multimanifold structure feature representation for quality prediction with an industrial application. IEEE Trans Ind Inf 2022;18(9):5849–58. http://dx.doi.org/10.1109/TII.2021.3130411.

[54] Jiang Y, Yin S, Dong J, Kaynak O. A review on soft sensors for monitoring, control, and optimization of industrial processes. IEEE Sens J 2021;21(11):12868–81. http://dx.doi.org/10.1109/JSEN.2020.3033153.

[55] Sun Q, Ge Z. Gated stacked target-related autoencoder: A novel deep feature extraction and layerwise ensemble method for industrial soft sensor application. IEEE Trans Cybern 2022;52(5):3457–68. http://dx.doi.org/10.1109/TCYB.2020.3010331.

[56] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: Teh YW, Titterington M, editors. Proceedings of the thirteenth international conference on artificial intelligence and statistics. Proceedings of machine learning research, vol. 9, Chia Laguna Resort, Sardinia, Italy: PMLR; 2010, p. 249–56, URL https://proceedings.mlr.press/v9/glorot10a.html.

[57] Zhu W, Mei B, Yan G, Ke Y. Measurement error analysis and accuracy enhancement of 2D vision system for robotic drilling. Robot Comput-Integr Manuf 2014;30(2):160–71. http://dx.doi.org/10.1016/j.rcim.2013.09.014, URL https://www.sciencedirect.com/science/article/pii/S0736584513000732.

[58] Romero-Ramirez FJ, Muñoz-Salinas R, Medina-Carnicer R. Speeded up detection of squared fiducial markers. Image Vis Comput 2018;76:38–47. http://dx.doi.org/10.1016/j.imavis.2018.05.004, URL https://www.sciencedirect.com/science/article/pii/S0262885618300799.

[59] Akinlar C, Tonal C. EDCircles: Real-time circle detection by edge drawing (ED). In: 2012 IEEE international conference on acoustics, speech and signal processing. 2012, p. 1309–12. http://dx.doi.org/10.1109/ICASSP.2012.6288130.

[60] Kuffner J, LaValle S. RRT-connect: An efficient approach to single-query path planning. In: Proceedings 2000 iCRA. millennium conference. IEEE international conference on robotics and automation. Symposia proceedings (cat. no.00CH37065). 2000, p. 995–1001. http://dx.doi.org/10.1109/ROBOT.2000.844730.

[61] Sucan IA, Moll M, Kavraki LE. The open motion planning library. IEEE Robot Autom Mag 2012;19(4):72–82. http://dx.doi.org/10.1109/MRA.2012.2205651.

[62] Mei B, Zhu W, Yan G, Ke Y. A new elliptic contour extraction method for reference hole detection in robotic drilling. Pattern Anal Appl 2015;18(3):695–712. http://dx.doi.org/10.1007/s10044-014-0394-6.