

RESEARCH ARTICLE

Novel model-based force and object slip estimation approach for neuromorphic vision tactile sensors

Murana Awad¹, Musa Omar Abdalla² , Mohammad I. Awad³, Yahya Zweiri¹ and Kinda Khalaf¹

¹Khalifa University, United Arab Emirates

²The University of Jordan School of Engineering, Amman, Jordan

³Abu Dhabi University, United Arab Emirates

Corresponding author: Musa Omar Abdalla; Email: admin@mechatronix.us

Received: 18 April 2025; **Revised:** 16 September 2025; **Accepted:** 31 December 2025;

First published online: 6 February 2026

Keywords: neuromorphic; vision tactile sensor; sensor modeling; event camera; LSTM classifier; force and slip detection

Abstract

Neuromorphic vision-based robotic tactile sensors fuse touch and vision, enabling manipulators to efficiently grip and identify objects. Precise robotic manipulation requires early detection of slips on the grasped object, which is crucial for maintaining grip stability and safety. Modern closed-loop feedback technologies use measurements from neuromorphic vision-based tactile sensors to control and prevent object slippage. Unfortunately, most of these sensors measure and report data-based rather than model-based information, resulting in less efficient control capabilities. This work proposes physical and mathematical modeling of an in-house-developed neuromorphic vision-based robotic tactile sensor that utilizes a protruded marker design to demonstrate the model-based approach. This sensor is mounted on the UR10 robotic manipulator, enabling manipulation tasks such as approaching, pressing, and slipping. The neuromorphic vision-based robotic tactile sensor-derived mathematical model revealed first-order system behavior for three manipulation-related actions under study. Experimental robotic manipulator grasping work is conducted to verify and validate the sensor's derived mathematical FOS model. Two data analysis approaches, temporal and spatial-temporal model based, are adopted to classify the manipulator-sensor actions. A long short-term memory (LSTM) temporal classifier is engineered to exploit the sensor's derived model. Also, the LSTM spatial-temporal classifier is designed using an event-weighted centroid of the region-of-interest features. Both LSTM methods successfully identified the robotic actions performed with an accuracy of more than 99%. Additionally, quantitative slip rate estimation is carried out based on centroid estimation, and qualitative assessment of pressing force is performed using a fuzzy logic classifier.

1. Introduction

Human hand abilities for manipulating carried objects and tools in unstructured environments are of great interest to scientists and engineers. The real-time dexterity and perception required to control objects and tools blindly make it even more challenging and beyond imagination. Despite the current level of advancement in the creation of dexterous grippers and human-like fingers by robotics manipulators, grasped objects and tools still require more stable and precise grasping. Improvements in tactile sensing and perception allow us to address such challenges with advanced feedback control strategies, enabling adjustments to grip force. However, one of the main difficulties that crucially hinders the development of superior tactile sensing is detecting an object's slip early. In the literature, numerous engineered tactile sensors have been proposed to approximate human abilities in tactile sensing and slip perception [1–3].

The slip phenomenon in robotics occurs when external forces acting on an object (e.g., weight and dynamics) exceed the frictional gripping force at the contact point, causing relative displacement

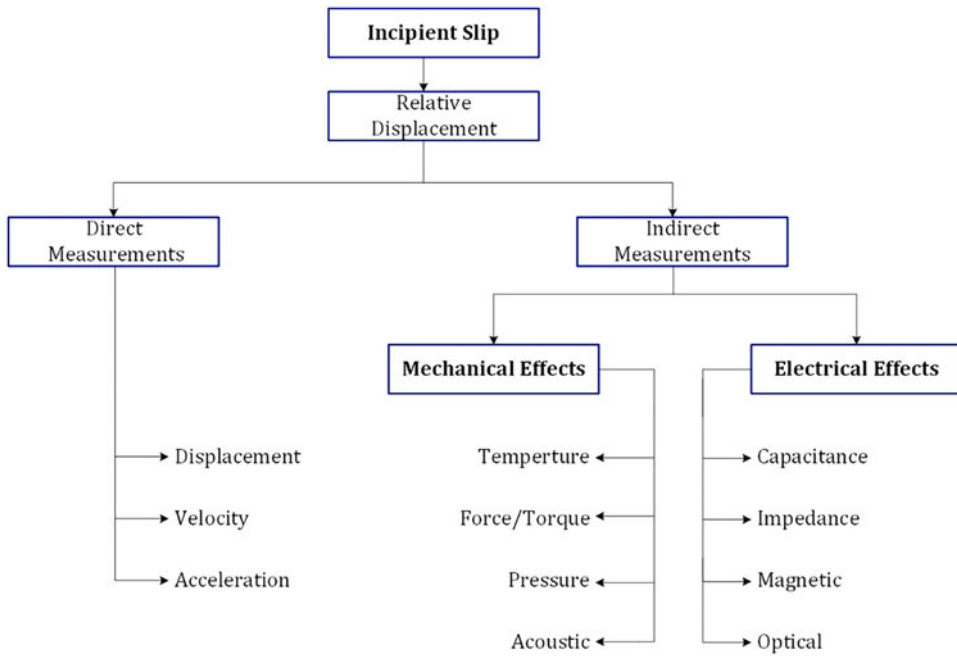


Figure 1. Current slip detection popular methods.

between the gripper and the object. This slip phenomenon, causes relative displacements, is a critical area of interest within the fields of robotics, grasping, instrumentation, and measurement. Such fields focus on detecting, measuring, monitoring, and controlling slips. These relative displacements are necessary conditions for slips and can be detected and quantified through direct and indirect measurement techniques.

Figure 1 presents a categorization of the popular relative-displacement-based techniques reported in the literature. Direct displacement measurements include displacement, velocity, and acceleration, while indirect measurements focus on detecting mechanical or electrical changes. Mechanical changes involve variations in temperature due to friction, force or torque, pressure, and acoustic waves at the contact surfaces. On the other hand, electrical changes encompass variations in capacitance, impedance, magnetic field, and optical properties induced by the contact surfaces [2–8].

Some robotic manipulator gripper designs are equipped with optical tactile sensors (OTS) to sense the conditions of the objects being grasped. OTS technologies utilize variations in light properties to detect object displacements, including reflection, refraction, diffraction, interference, dispersion, and polarization. Generally, OTSs utilize optical receptors to detect variations in specific light properties emitted from a light source or transmitter toward a tactile medium. These optical receptors can be single modules or arranged in arrays, similar to those found in cameras. Depending on applications and requirements, vision-based OTS engineering designs employ frame and event cameras. On the other hand, well established engineered-shaped tactile medium materials are used to interact directly with the object during grasping. These tactile materials are designed to harness the light properties of optical receptors, such as markers, patterns, flexible membranes, and guided waves [6, 9–11].

Many novel and popular neuromorphic vision-based tactile sensors have been developed and extensively studied for their attractive capabilities in fast perception, slip detection, and object manipulation. Such sensors include, but are not limited to, GelSight [12], TacTip [13], NeuroTac [14], Tactigraph [15], Kumagai sensor [16], FingerVision [17], and Evetac [10].

Currently, many vision-based tactile sensors are limited to detecting incipient slip without quantifying the amount of slip displacement or the slip rate for grasped objects. Real-time active grasping force

control requires slip displacement estimation or the degree of slip as feedback signals to adjust the pressing force on the object. In the literature, three popular strategies are used to estimate the degree of object slips in tactile sensors: the stick/slip region-based method, the friction estimation-based method, and the micro-element resultant (MER) force-based incipient slip degree evaluation method [18]. These strategies are based on contact mechanics approaches [19] and depend heavily on the object's material type and properties.

Vision-based robotic tactile sensors operate in data-based, model-based, or hybrid modes. In model-based tactile sensors, tactile sensors use mathematically generated models (i.e., using physical and engineering principles) to map or interpret sensor input readings (e.g., from a camera or force sensors) into information about the object in contact, such as press and slip, force distribution, shape, or texture. In contrast, data-driven methods rely on extensive multi-scenario training datasets. In particular, model-based tactile sensors with event cameras provide a mathematical framework for reconstructing contact features by mapping image intensities or deformation data to contact-related phenomena, such as object slip. Both approaches can be used in Artificial Intelligence (AI), but the model-based approach appears to have advantages over the data-based approach. Flexibility, higher accuracy, reduced reliance on extensive training data, and compatibility with controller design are among the significant benefits offered by the model-based approach [20].

This work presents a novel model-based approach to vision-based robotic tactile sensors. A newly developed in-house, novel neuromorphic event-based tactile sensor demonstrates how a model-based approach can provide fast, rich output and AI-enabled data. The sensor's ability to estimate slip displacement and the degree of pressing force on grasped objects gives it an advantage over many other neuromorphic event-based tactile sensors. Furthermore, understanding the sensor's mathematical model facilitates comprehension of its operations and limitations, particularly during a robotic manipulator's grasping for an action event. Furthermore, this model will provide a framework for implementing predictive and adaptive control-based models to improve the forecasting of robotic grasping strategies.

Finally, the proposed model-based approach is verified experimentally using an in-house-developed neuromorphic tactile sensor setup on an industrial robotic manipulator to demonstrate its ability to detect and quantify the objects translational slips and the degree of pressing force at an early stage. Such a slip-displacement estimate, which is needed as a feedback signal for most controllers, is computed as the weighted centroid of the sensor's marker strip region of interest (ROI). Although many researchers used the marker's centroid feature for translational and rotational slip detections, it was estimated only within a force field [9, 21, 22]. In contrast, the proposed tactile sensor does not rely on a force field generated by the markers; instead, it uses variations in the markers' strip aperture areas. During pressing or slip actions, the strip deformation results in variations in the aperture area (i.e., the trackable feature) and enables centroid variations that are easily mapped to slip-rate estimates, in contrast to full n-dimensional field computations.

The novelty of the used sensor lies in its bioinspired, flexible-skin design, featuring protruding markers that mimic the ridges of human fingerprints, which are associated with heightened tactile sensitivity in humans. Unlike conventional tactile sensors, which commonly use flat or internally embedded markers, our sensor's protruding markers enable detection of low contact forces and incipient slip. Furthermore, these markers are integrated with an event-based neuromorphic vision sensor that captures their deformations. Unlike frame-based cameras, this enables high temporal resolution and low-latency detection of tactile events such as pressing and slipping. Finally, the sensor's flexible skin is fabricated using readily accessible, low-cost 3D-printing technology (TPU: Ninjabflex). This fabrication method reduces complexity and cost while enabling rapid customization and prototyping.

This paper's layout begins with an overview of slip theory. Then, it introduces the novel, in-house-designed neuromorphic vision-based tactile sensor with its event camera. A mathematical model of the sensor is derived and verified experimentally using a robotic manipulator. Finally, experimental results, discussions, and conclusions are presented.

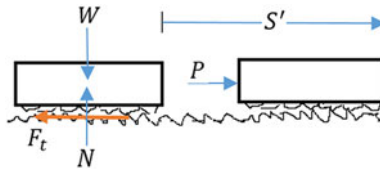


Figure 2. Two bodies' magnified surface.

2. What is slip?

Classical mechanics models slip using Coulomb's principle, which states that frictional forces are proportional to normal forces. To illustrate the slip/slide phenomenon from the physical point of view, consider a body slipping or sliding over another body surface in the presence of friction, which is depicted in Figure 2. It should be clear that even for rigid bodies, the contact area is deformable on the microscopic level with consequential effects for the teeth in the contact area to experience adhesion, breaking-off points, pull-away effects, and interlocking effects with vibration. On the other hand, additional effects, such as stiction, may occur for deformable objects. Nevertheless, slipping and sliding displacement occurs when the applied force exceeds the tangential static frictional force.

When a constant applied force P generates enough work to displace an object and overcome the resultant tangential frictional force F_t , it accelerates the body until it reaches a constant velocity v , (i.e., dynamical equilibrium occurs at t_1). Taking the energy balance between two time instances t_1 and t_2 , where $t_2 > t_1$ gives $W_{force} - W_{friction} = \Delta KE$, which may be used to model all these effects as shown by the following equation:

$$\frac{1}{2}mv_1^2 + PS - F_t S' = \frac{1}{2}mv_2^2 \quad (1)$$

where $F_t = \mu_k N$ is the frictional tangential force, μ_k is the kinetic friction coefficient, N is the normal force, P is the constant external applied force, $v_1 = v_2 = v$ for a constant speed, S is the planned theoretical object displacement, while S' is the actual or tangible displacement of an object.

Finally, according to Coulomb's principle, the tangential frictional forces remain constant during the object's displacement. But depending on the object's surface finish quality, some of the object's Kinetic energy is transformed into other forms of internal energy, $\mu_k N(S - S')$, such as thermal, vibration, and acoustic. Consequently, these internal energy effects enabled researchers and engineers to use indirect measurement techniques for object slips in robotics applications [23–25]. One popular modeling technique of these frictional effects is to use an internal equivalent damping effect that is proportional to the slip change rate (i.e., $v(t)$), that is $F_t(t) = C dS'/dt$, where C is an equivalent damping coefficient. However, this work considers only the actual, tangible (i.e., directly measurable) displacement as a measure of slip quantification.

In the literature, some researchers model incipient slips by monitoring changes in the object-gripper contact area [5]; however, this approach does not provide useful information on displacement slips or slip rates. In contrast, the energy-based modeling of slip, as shown in Eq. (1), is favored.

3. Neuromorphic model-based tactile sensors

Neuromorphic vision-based tactile sensors offer spatial, temporal, and AI-enabled advantages for a diverse range of robotic grasping applications. Subsequent subsections present an overview of these sensor types, and an in-house-designed neuromorphic vision-based tactile sensor is introduced to illustrate a novel model-based approach that reduces reliance on massive training data. Unlike data-driven methods, model-based approaches can be effective without requiring extensive training datasets and offer the potential for high accuracy and flexibility. Physical models can accurately represent deformation and force distribution and could be adapted to different sensor types and applications. Finally, the model-based approach offers better control compatibility by enabling model-based controller design, which is far superior to modelless controller design approaches [21].

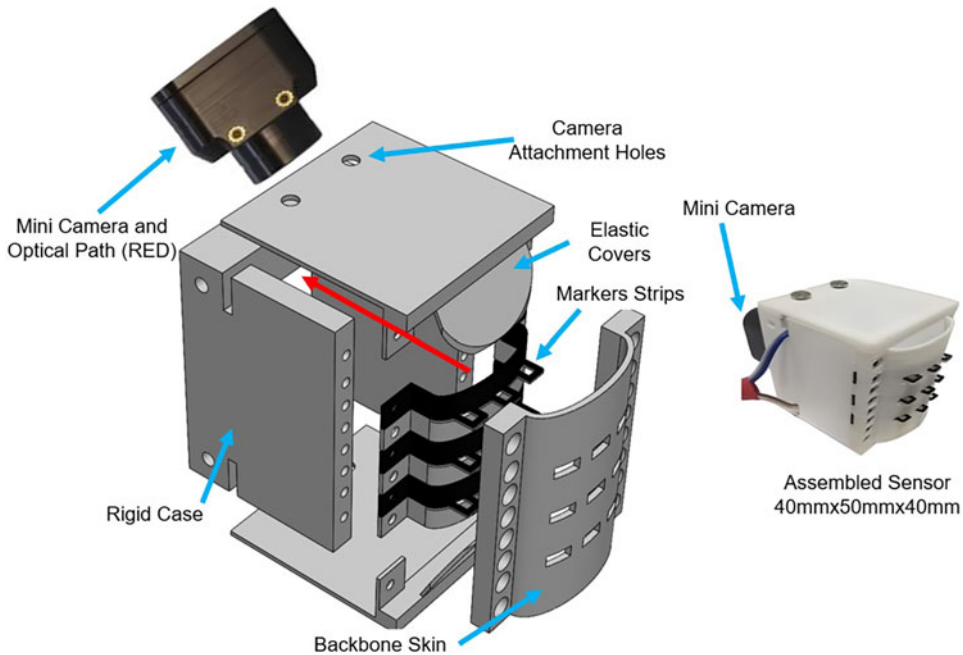


Figure 3. *Novel neuromorphic dynamic vision sensor.*

3.1. Neuromorphic dynamic vision sensors

Recently developed neuromorphic event-based dynamic vision sensors (DVS) use event cameras with patented technology that mimics the human retinal response. First, light is reflected from surrounding objects and enters through the cornea (i.e., the aperture), pupil, and lens, ultimately hitting the retina. The photoreceptor cells in the retina respond to light, converting its energy into an electrical signal. Furthermore, this signal travels through the human optic nerve to the brain, where it is finally converted to portray the images of objects [26, 27].

Neuromorphic DVS integration of the recently developed cameras enables them to capture detailed visual information about objects' spatial properties, dynamics, and environmental interactions. Both frame and event cameras are used in designing state-of-the-art neuromorphic DVS for robotic grasping applications [28, 29].

Neuromorphic event-based cameras generate asynchronous events triggered by significant changes in brightness intensity at individual pixels. This on-demand operation offers advantages in the design of robotics tactile sensors over frame-based cameras, such as high temporal resolution, low energy consumption, fast response, and a broad dynamic range [30].

3.2. A novel-designed neuromorphic dynamic vision sensor

Figure 3 illustrates our in-house, bio-inspired, neuromorphic vision-based sensor that uses event-based vision technology to rapidly detect tactile interactions. This novel neuromorphic sensor is designed to emulate the high tactile sensitivity of human fingerprint ridges, featuring protruding markers that detect the slightest touch or minimal contact. These markers, alongside the sensor's backbone skin, are printed using Ninjabflex material. The sensor's entire assembly is encased in a PLA-based 3D-printed housing, and internally, a DVXplorer Mini neuromorphic event camera is secured to capture real-time, dynamic deformations of the sensor's markers and strips under various contact scenarios. The sensor's actual dimensions are $40 \times 50 \times 40$ mm. The complete design and fabrication of this sensor are described in ref. [30].



Figure 4. Event camera change in intensity.

This sensor is designed to detect early slips in robotic grasping applications. Embedding this sensor within a robotic manipulator's end-effector will measure data on pressing and slipping with rigid and moderately soft objects. During initial grasping, the sensor's markers will be partially axially pushed, causing changes in brightness through the aperture holes and giving data about the pressing operation. If the grasped object starts to slip, causing relative displacement, the sensors' marker strips will bend proportionally, altering the aperture area and brightness (i.e., marker strips). The red arrow in Figure 3 illustrates the optical path for the tactile sensor, starting from the external light source, being reflected through the aperture area of the marker strips, and ending at the event camera as a change in intensity.

Figure 4 illustrates the generated light diffractions through the aperture and the detection by the DVXplorer Mini event camera [31]. The used event camera reports time-stamped relative logarithmic changes in the light intensity on the pixel level in the form of occurring events $\{+1, -1\}$ based on predefined ON/OFF thresholds. Positive indicates an increase, while negative suggests a decrease in logarithmic intensity, and no changes in logarithmic intensity will be reported.

$$Event = \text{sgn} \left(\log \left(\frac{I(x, y, t + \Delta t)}{I(x, y, t)} \right) \right) \quad (2)$$

where $Event \in \{+1, -1\}$ are the reported events of a pixel that is located at (x, y) , $I(x, y)$ is the pixel's intensity states, which occurred at a time t and at the time $t + \Delta t$, respectively.

Event cameras do not send their captured high-definition images at a fixed frame rate, unlike conventional frame-based cameras. Instead, only events that occur when movement or change happens in the scene are produced, significantly reducing power consumption, data rate, and computational requirements. Even ultrafast cameras are designed for demanding applications in robotics, motion analysis, aerospace, AR/VR, and autonomous vehicles [32].

A DVS-generated event is documented using the address event representation (AER) format: $E = [x, y, t, p]$, which records the pixel address (x, y) , the time at which the address was reported, t , and the polarity of the brightness change (i.e., increase or decrease).

For the proposed model-based approach, we must establish a mathematical relationship between the event camera parameters and the sensor used in this work. In optometry, the viewable area that the camera lens can capture is referred to as the field of view (FOV), while the strength of the lens's focus (i.e., convergence) or the divergence of light is referred to as the focal length (f). The FOV determines the portion of the object that is cast onto the camera's sensor. This enables correlating the object's physical area, which can be imaged and modeled as a horizontal or vertical FOV in millimeters, or as an angular FOV specified in degrees. Referring to Figure 5 and using mathematical geometry, the following equation relates the sensor's marker strips aperture area to the camera sensor array's vertical size [32]:

$$H = \frac{f}{\delta} FOV \quad (3)$$

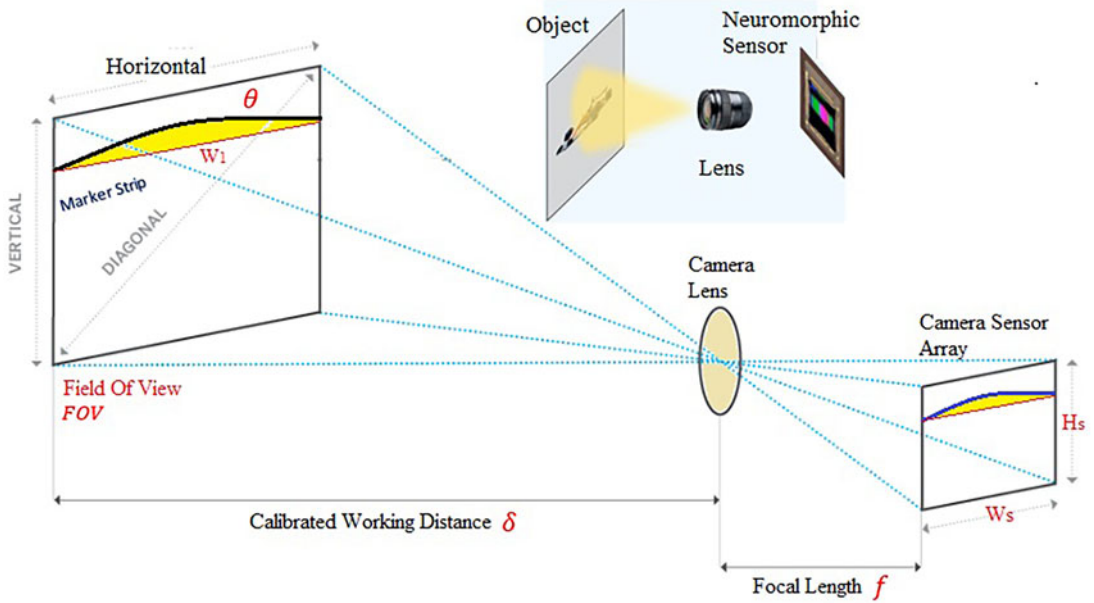


Figure 5. Object and captured image correlation.

$$A_p = l_2^2 \theta \quad (4)$$

where H is the camera sensor vertical size, f is the used camera focal length, and 3.6 mm lens, δ is the strip object setup distance that is the distance from the strip object to the camera after calibration, FOV is the strip object's FOV, A_p is the marker strip opening area, l_2 is the sensor dome radius, and θ is the marker strip angular displacement. Figure 5 illustrates the overall light optical path, the event camera parameters, and the relationship between FOV , sensor size H , and δ for a given angular FOV ($AFOV$).

3.3. The proposed neuromorphic DVS sensor modeling

The sensor's markers and strips are made of Ninjaflex, a material that is rubber-like in its flexibility and damping. To simplify the model without compromising accuracy, the strips will be modeled using spring and damper elements, as shown in Figure 6. The markers pressing action will be physically modeled using linear stiffness and damping, represented by $K1$ and $B1$, respectively. In contrast, the fixed-end marker strips will undergo rotational motion during the slip action, which will be physically modeled using linear and rotational stiffness and damping coefficients, $K2$ and $B2$, respectively. Finally, the inertial effects of the markers and strips will be neglected due to their small masses and low slipping rates.

In robotic applications, objects are first secured in a stable grip but may suffer undesired slips due to dynamic effects during motion, which can propagate to a complete object slip. The initial gripping action will push the sensor's markers axially, causing a small angular displacement θ_0 of the markers' strips. However, a slip will cause the markers to bend and cause further angular displacement of the markers' strips that is proportional to the object's slipping displacement.

As mentioned earlier, the mathematical modeling of the sensor's actions, pressing and slipping, will be treated in two steps. Referring to Figures 6 and 7, first consider a robotic action that requires an external force P , which may be decomposed into two force components P_1 and P_2 . Consequently, the dynamical equilibrium is given by

$$\sum F_i(t) = m_s \ddot{x}(t) = 0 \quad (5)$$

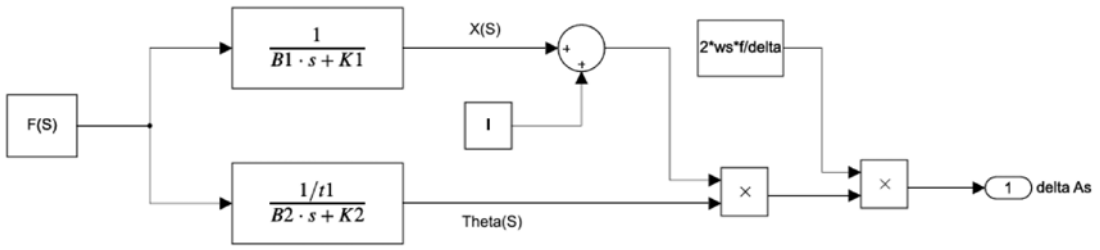


Figure 8. Sensor block diagram equations summary.

The sensor's markers undergo two types of force-generated displacements: axial and bending of the marker strips. Figure 7 illustrates the displacements of bending marker strips (i.e., those on the light optical path) modeled as cantilevered beams. Using basic beam theory, the beam's bending profile (i.e., elastica curve) may be correlated to the aperture area, A .

The length of the deflected part of the beam l_b varies depending on the degree of pressing axial force P_1 on the grasped object. On the other hand, the lateral force P_2 causes the marker to bend, and the maximum deflection y_{max} is given from the cantilevered beam theory as

$$y_{max} = \frac{P_2 l_b^3}{3EI} = \frac{P_2 (l - x_c)^3}{3EI} \quad (12)$$

$$P_1 = K_1 x_c \quad (13)$$

where x_c is the beam end point 'C' displacement, l_b is the beam's cantilevered length, l is the beam's total length, E is the material modulus of elasticity of the marker, and I is the moment of inertia for the marker's rectangular cross-section. Consequently, the marker object slipping force P_2 will cause the marker strip to bend and vary the aperture area in front of the event camera.

$$\Delta A_p = l_2^2 \Delta \theta = \frac{l_2^2}{l_b} y_{max} \quad (14)$$

$$\Delta A_s = w_s \Delta H = w_s \frac{f}{\delta} \Delta FOV = w_s \frac{2f}{\delta} \Delta y_{max} \quad (15)$$

where ΔA_p is the marker strip, a small change in the aperture area, ΔH is the corresponding affected vertical pixel array in the camera sensor, and w_s is the fixed sensor marker strip aperture window width. The pixel array has a resolution of 640×480 and measures $640 \text{ pixels} \times 9 \mu\text{m/pixel} = 5.76 \text{ mm}$ wide and $480 \text{ pixels} \times 9 \mu\text{m/pixel} = 4.32 \text{ mm}$ high. Figure 8 illustrates the integration of all previous equations in block diagram form, where an input pressing force results in a change in aperture area on the sensor, which is recorded by the event camera.

4. Experimental work

4.1. Experiment description

The following laboratory experiment demonstrates the operation and testing of a novel model-based neuromorphic dynamic vision sensor. Figure 9 illustrates the integration of the sensor with an end-effector on a UR10 Robotic Manipulator. The robot's end-effector and sensor will press against a dummy target plate to simulate a gripped object. The selected target plate is rigid, flat, and has a smooth surface finish, fabricated from PLA material using 3D printing techniques. The robot's end-effector will move at a constant speed to simulate a wide range of tactile interactions with its slips. The interaction starts by moving the robot from a start home position and placing the sensor at an arbitrary reference position close to the target plate (i.e., approaching). Then, the robot manipulator is instructed to perform the pressing by advancing the robot manipulator's prespecified displacement amounts. Finally, the robot will move vertically, mimicking a slip for a predefined displacement, then release the press and return

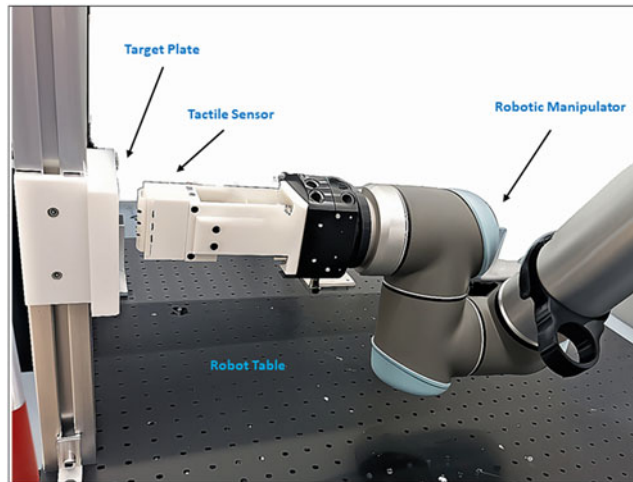


Figure 9. Experimental tactile sensor setup.

to the home position. Performing all experiments from a home position with constant robot speed and selecting prespecified displacements guarantees repeatability of the pressing force experiments, which is given as $P = B \frac{dx}{dt} + K \Delta x$, function of speed and displacement.

The experiment design is engineered to provide different levels of applied pressing forces against the target plate. The selected force levels are labeled as light, small, small medium, medium, small, big, and big pressing levels. Four experimental trial groups were carried out, and each trial consisted of twelve different extents of applied pressing forces on the sensor markers.

Finally, the robot operating system establishes as accurate and synchronized experimental datasets as possible. It enabled the control of the UR10 robot to precisely execute robotic actions such as approaching, pressing, sliding, and releasing while simultaneously managing data acquisition from the DVXplorer event-based camera. This ensures seamless synchronization between the robot's movements and sensor data, providing accurate and repeatable experimental conditions.

4.2. Experimental results modeling

Three sets of raw data generated by neuromorphic tactile sensors are captured and depicted at the top of Figure 10. The three curves represent three degrees of applied force: small, medium, and big. A 20 ms of accumulated raw events are used to identify four distinct regions corresponding to the robotic manipulator's actions. The first region represents the robotic manipulator as it approaches the target plate. The second region represents a pressing action against the target object, while the third region shows the object slipping action. Finally, the last region represents the object's release action. The bottom graph in Figure 10 illustrates the same data in a filtered form, where moving averages are used to filter and smooth the data, allowing the four action regions to be clearly identified.

The regions of pressing and slipping behavior portrayed in Figure 10 clearly validate the first-order system (FOS) derived mathematically in Eqs. (10) and (11). Figure 11 depicts two grasping axial pressing force contact levels applied on the sensor, "low" and "big" contacts. The left graph shows the low-contact pulse-pressing axial force response, which represents a typical FOS behavior for the pressing action. In contrast, due to the nonlinear behavior of the Ninja flex material, the right-hand graph's "big" contact force response exhibits a higher level of nonlinear, unmodeled dynamics. Considering only the linear FOS behavior and using an exponential unit pulse input, the FOS curve fitting may be given as follows:

$$x(t) = \begin{cases} \alpha_1(1 - e^{-t/\tau_1}) & 0 \leq t < T \\ \alpha_1(e^{-(t-T)/\tau_1} - e^{-t/\tau_1}) & t \geq T \end{cases} \quad (16)$$

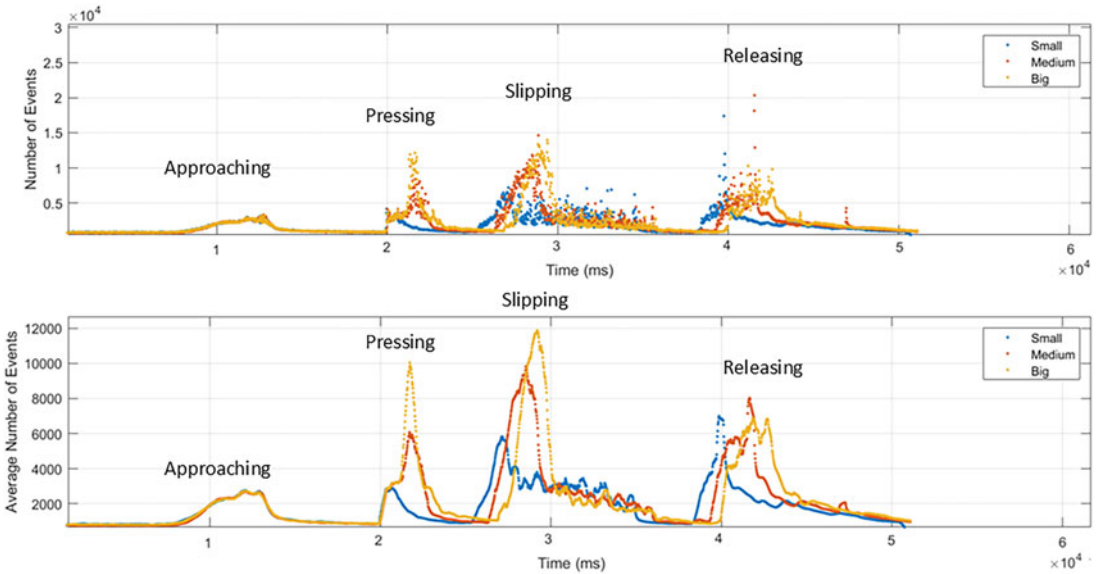


Figure 10. Three datasets of event tactile sensor raw data visualization.

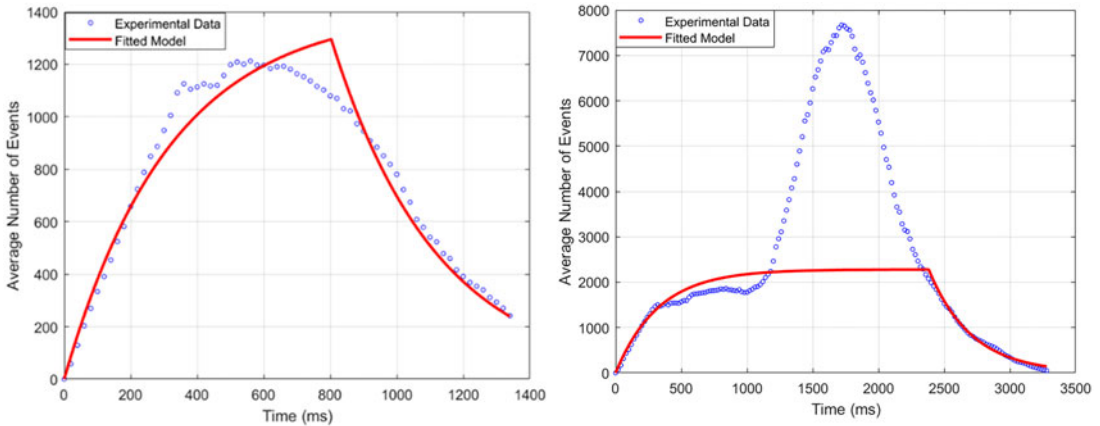


Figure 11. Unit step input force model fitting for the pressing FOS.

where $\alpha_1 \in [1365, 2280]$ and $\tau_1 \approx 300$ ms. The red curves in Figure 11 represent the fitted FOS linear model with a Pearson correlation coefficient of $R = 0.983138$, which was used as training data for the AI classifier.

Figure 12 shows the slip action of the experimental and fitted models for a unit-ramp input. The FOS response may be captured using the following equation:

$$\theta(t) = \alpha_2 \left(\frac{1}{\tau_2} t - 1 + e^{-t/\tau_2} \right) \quad (17)$$

where $\alpha_2 \approx 1615$ and $\tau_2 \approx 325$ ms. The FOS model is fitted with a correlation coefficient of $R = 0.999762$.

4.3. Experimental results visualization

The sensor actions' experimental data may be visualized as an image. Spatially visual-mapped camera frames are constructed by mapping event coordinates onto a 2D grid corresponding to the sensor's camera resolution of 480 by 640. The frames are segmented by dividing the event stream into frames

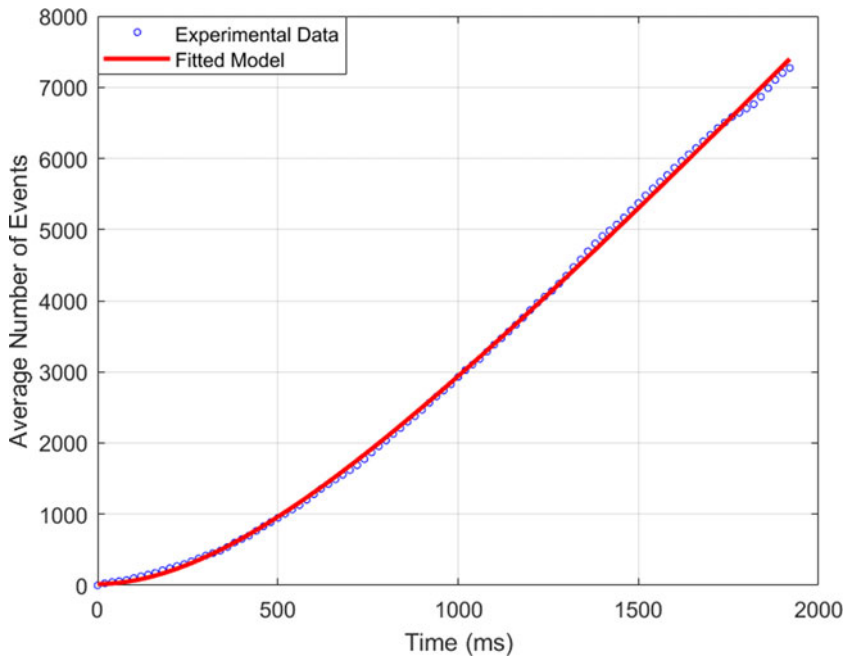


Figure 12. Unit ramp input force model fitting for the slipping FOS.

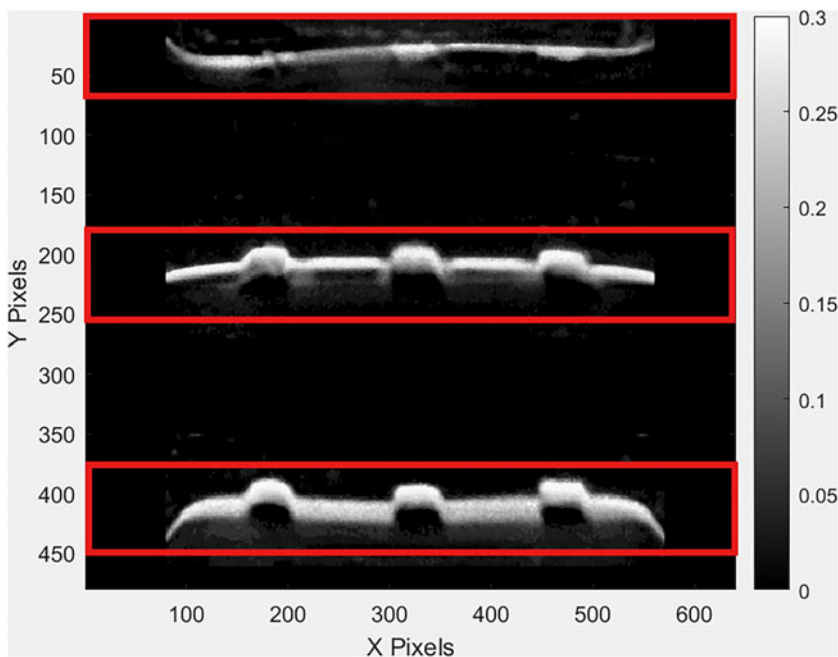


Figure 13. Slip frame picture visualization, deformed markers.

based on an accumulation time delta. Furthermore, noise suppression and reduction in accumulated frames (i.e., salt-and-pepper noise) were achieved using a median filter.

Figure 13 shows the picture frame, which displays the sensor's markers and strip displacement during a slip action. These highlighted rectangular areas are used as a ROI in subsequent analysis.

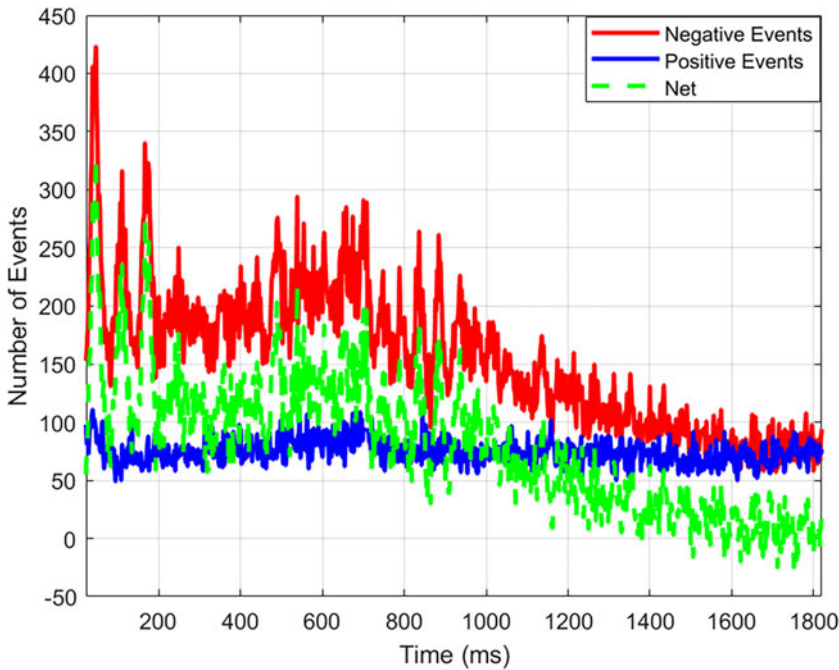


Figure 14. Pressing events polarity issues example.

4.4. Experimental pressing raw results

The raw data events comprise positive and negative polarities; positive marks increase in event intensity, while negative marks represent a decrease in event intensity. Consider a medium-pressure force action, as shown in Figure 14, which exhibits this dual behavior. During the pressing action, two interrelated changes in intensity occur. The sensor markers block ambient light and lower their intensity as they are pressed. On the other hand, the pressing action opens the marker strips further, resulting in increased intensities in the camera's vicinity. To decouple these interrelated effects, one can find the difference between these events, shown in the blue curve.

Furthermore, quantifying the amount of pressing force (i.e., “light,” “small,” “small medium,” “medium,” “small big,” and “big”) is characterized by the increased number of sensor-reported events. One way to quantify the pressing amount is to use the area under the designated pressing region as portrayed in Figure 10. Consequently, Figure 15 depicts the strong correlation between the degree of pressing and the computed area in the selected region.

5. Data analysis

The classification analysis stage starts after collecting, conditioning, and accumulating the sensor's data. This work considers two types of sensor event data classifications: Temporal and spatial-temporal long short-term memory (LSTM) classifications. Figure 16 is a flowchart diagram summarizing the data analysis process used to classify the robotic manipulator and sensor actions, namely, approaching, pressing, and slipping.

A recurrent neural network (RNN) is widely used in deep learning to classify events or actions qualitatively. It may be trained to recognize previously presented robotic manipulator actions depicted in Figure 10. Unfortunately, the nature of the experimentally collected data is spatial, temporal, and sequential, making it difficult for conventional RNN algorithms to process classifications over long-term dependencies. Another variation of RNN capable of processing long-term dependencies needed for sequential predictions is the LSTM network. LSTM incorporates feedback connections, enabling it

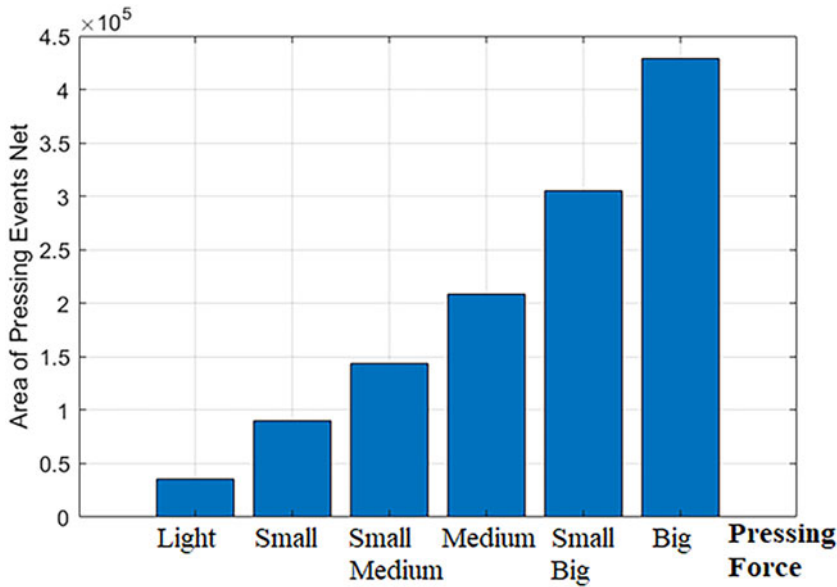


Figure 15. Pressing force qualitative extent.

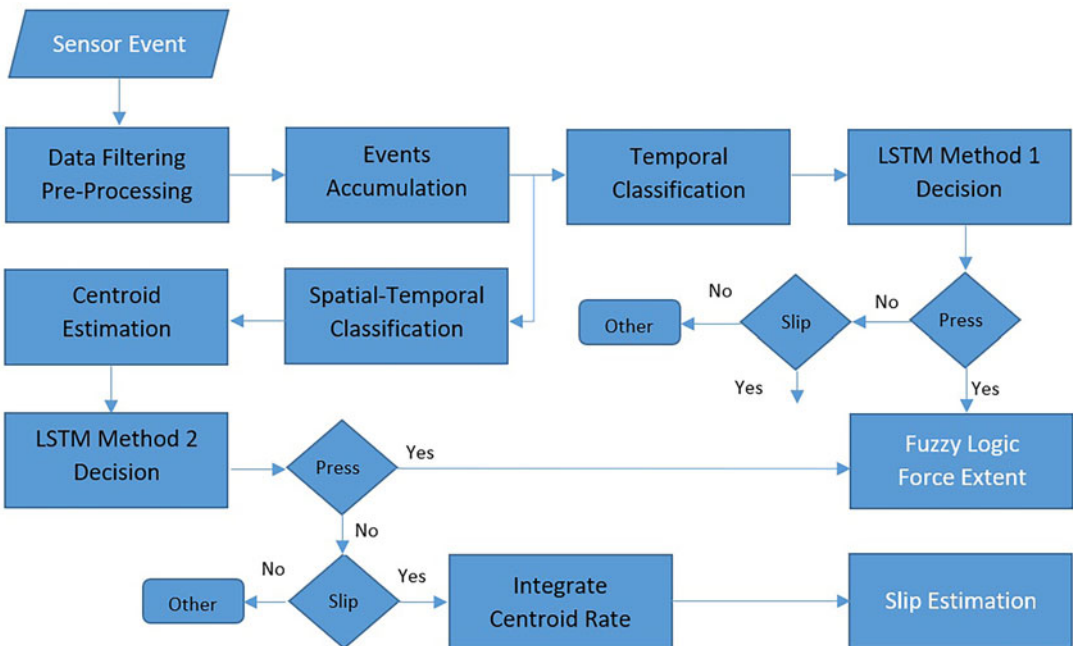


Figure 16. Data analysis proposed process.

to process entire data sequences rather than just individual data points. This makes it highly effective at understanding and predicting patterns in sequential data, such as time series. A complete description and review of the LSTM network operations may be viewed in reference to ref. [33].

The neural network architecture comprised five layers: an input layer, a single LSTM layer with hidden units to capture temporal dependencies, a fully connected layer, a Softmax layer, and a classification layer. Despite the architecture's simplicity, the LSTM layer effectively learned the sensor's FOS patterns from tactile sequences, enabling accurate classification of the approaching, pressing, and sliding actions.

Table I. Temporal LSTM classifications results.

Metric action	Precision	Recall	F1_Score	Overall test accuracy
Approaching	0.99684	1.0	0.99841	99.89%
Pressing	1.0	0.9998	0.99991	
Slipping	0.9998	0.99683	0.99833	

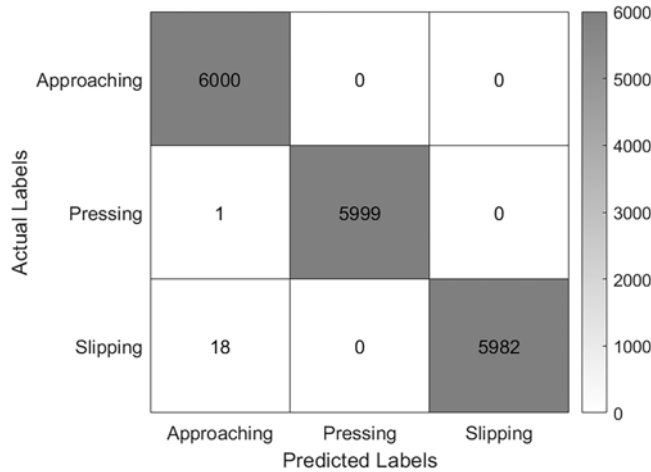


Figure 17. Temporal LSTM confusion matrix.

5.1. Temporal LSTM classification

Model-based LSTM classification can be utilized to leverage the previously identified FOS behavior. This work generated four experimental trials based on the derived model, which were used to train the LSTM neural networks. Each trial consisted of twelve applied pressing forces with different grasping force severities or extents. The experimentally identified FOS dynamical model is used to generate experimentally trained data for the spatial–temporal dynamical LSTM neural nets. The identified model is used to generate numerous FOS training trajectories for neural network training, using a diverse set of random robotic actions and their corresponding initial conditions.

Referring to Figure 16, the data analysis process diagram, the first step is to qualitatively classify the robot’s action type using a temporal LSTM network. The second step estimates the extent of the pressing force qualitatively using a fuzzy logic (FL) classifier. The LSTM results are shown in Table I, and the corresponding confusion matrix is shown in Figure 17.

5.2. Spatial–temporal LSTM classification

A model-based LSTM classification approach using the weighted centroid feature is applied similarly to the temporal LSTM training, with four model-based experimental trials generated to train the spatial–temporal LSTM neural nets. Each experimental trial consisted of twelve applied pressing forces with different grasping force severities or extents.

Referring to the Figure 16 data analysis process diagram, the first step is to qualitatively classify the robot’s action type using the spatial–temporal LSTM based on the weighted centroid estimations of the ROI. The second step estimates the extent of the pressing force qualitatively using a FL classifier.

Finally, the weighted centroid dynamics is used to calculate the slip rate and compute the traveled displacement in Δt time. The spatial–temporal LSTM classification results are shown in Table II, and the corresponding confusion matrix is shown in Figure 18.

Table II. Spatial–temporal LSTM classifications results.

Metric action	Precision	Recall	F1_Score	Overall test accuracy
Approaching	0.99605	0.99923	0.99764	99.4054%
Pressing	0.98301	0.98424	0.98363	
Slipping	0.99667	0.99347	0.99507	

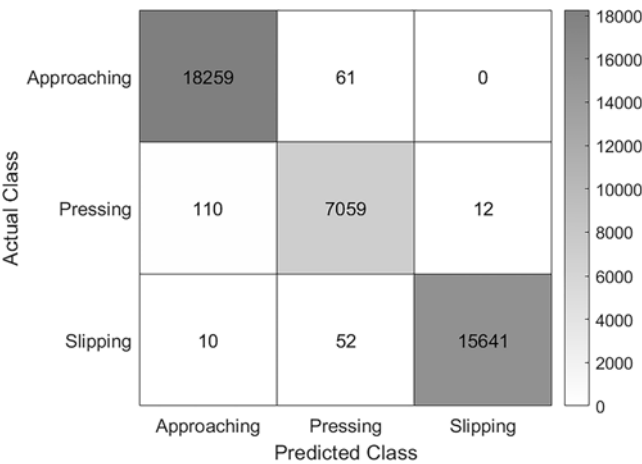


Figure 18. Spatial–temporal LSTM confusion matrix.

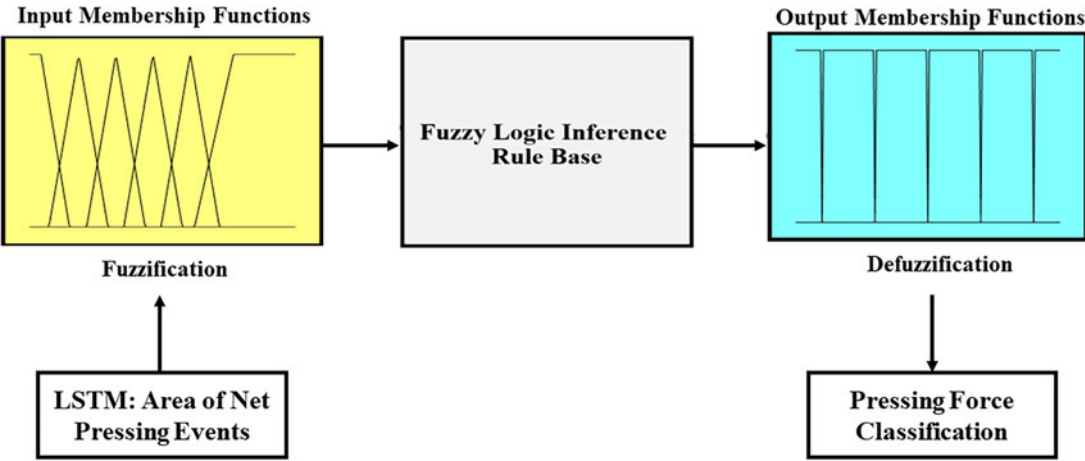


Figure 19. Fuzzy logic pressing force estimation.

5.3. Fuzzy logic pressing force estimation

The estimation of the extent of the pressing force is achieved using a FL-based classifier. The proposed FL classifier comprises four key steps: fuzzification of all crisp inputs, a rule base for classification rules, inference for decision-making, and defuzzification for the outputs. Figure 19 illustrates the structure of the proposed FL classifier [34].

The fuzzification and defuzzification of the input and output linguistic variables and their corresponding interval ranges are captured using the fuzzy membership functions illustrated in Figures 20 and 21, respectively. The overlaps between the membership functions are embedded to account for boundary values and induce continuous, smoother transitions, allowing for multiple rules to fire [35].

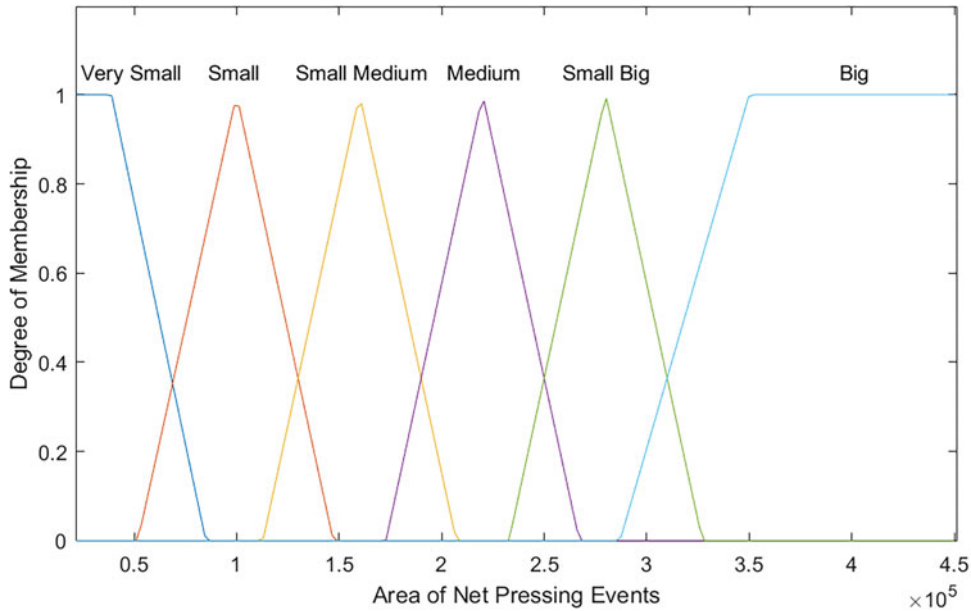


Figure 20. FL input fuzzification.

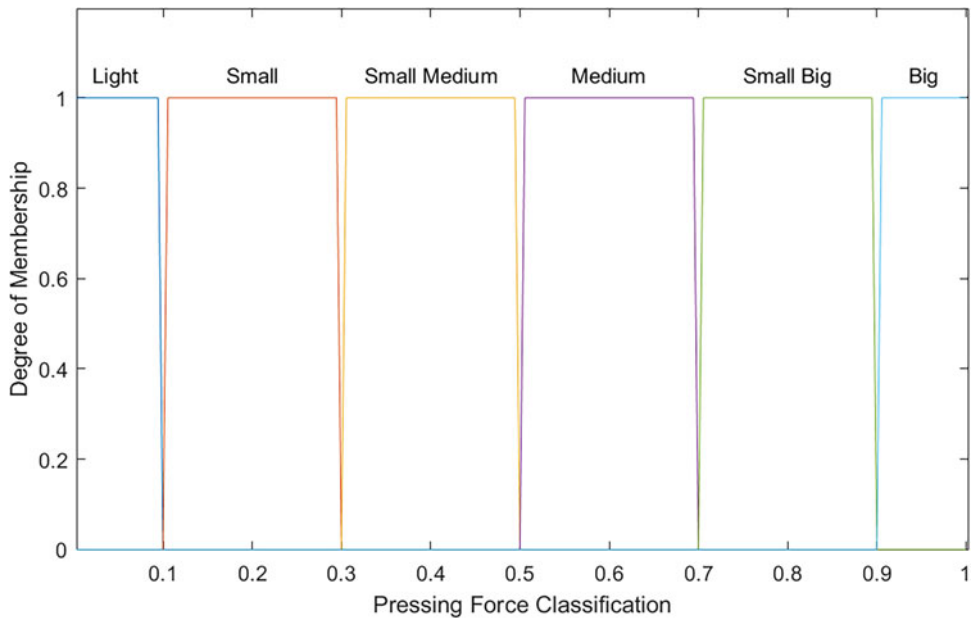


Figure 21. FL output fuzzification.

The pressing force classifier rule base is simple in structure yet effective. Table III summarizes the simple rules used in the classifier. Each rule is evaluated using the **If-Then** statement. To illustrate the operation of the FL classifier, Figure 22 is used for illustration. Considering a pressing force action classification, where two adjacent rules are fired simultaneously (i.e., due to input overlap), the corresponding output areas are aggregated to produce their centroid.

Ultimately, a thorough analysis reveals that the weighted centroid feature effectively represents events and regional concentrations. The weighted centroid of the ROI correlates with the object's slip action. Hence, by continuously evaluating the centroid feature over time, we can estimate the slip velocity

Table III. FL rule base.

IF (area of net pressing events)	Then (pressing force classification)
Very small	Light
Small	Small
Small medium	Small medium
Medium	Medium
Small big	Small big
Big	Big

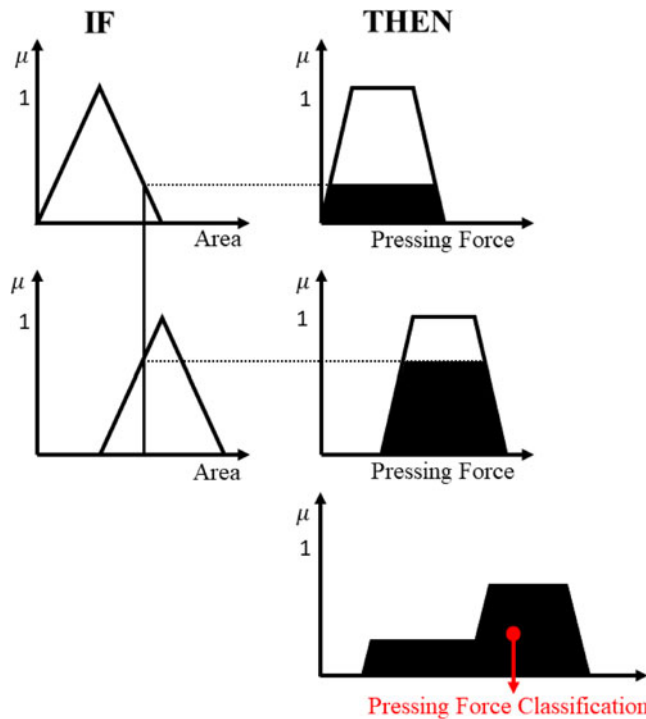


Figure 22. FL rules firing and inference example.

profile, enabling us to calculate the slip displacement at any instant. Spatial–temporal tracking of the centroids over time is used to analyze movement patterns within each *ROI*.

$$RoI = \{(x_i, y_j) : i, j \in \text{corners of } RoI\} \quad (18)$$

Using the events intensities as weights, we shall use a weighted centroid to account for the varying intensities within the *ROI*.

$$\bar{x} = \frac{\sum x I(x, y)}{\sum I(x, y)}, \bar{y} = \frac{\sum y I(x, y)}{\sum I(x, y)} \quad (19)$$

where $I(x, y)$ is the intensity at the pixel $(x, y) \in RoI$ after performing data filtering.

Finally, the object's slip displacement may be estimated from the centroid's rate of change by implementing numerical integration. The integration acts as a low-pass filter, reducing the effects of high-frequency noise.

$$\vec{v}(t) = \frac{1}{\Delta t} (\bar{x}(t)\hat{i} + \bar{y}(t)\hat{j}) \quad (20)$$

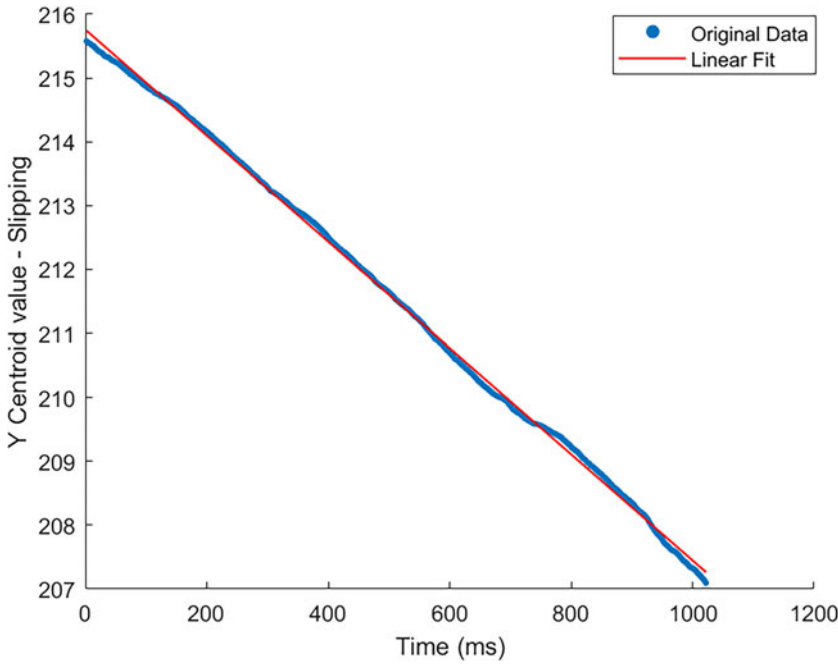


Figure 23. Centroid vertical trajectory.

$$\vec{d}(t) = \int_{t_1}^{t_2} \vec{v}(t) dt \quad (21)$$

The change in the centroid coordinates in the ROI ($\bar{x}$, $\bar{y}$) provides information about the object's slipping trajectory at its contact point. In this work, only translational information is reported by the tactile sensor. Consequently, only the $\bar{y}_{coord}$ is used to estimate the object's slip displacement. Figure 23 illustrates the linear behavior of the $\bar{y}_{coord}$ as the object is slipping for a medium grasping force extent, and the negative slope is from the orientation of the pixel grid (i.e., during the slip, the object falls while markers move up) as shown in Figure 13. The slope of the linear regression line should represent the speed of the slip, which reflects the robotic manipulator end-effector speed. The estimated average slope for the experimental trials is approximately 0.005 m/s, while the robotic manipulator end-effector speed is 0.004 m/s. This demonstrates the sensor's ability to estimate the object's slipping speed and displacement.

6. Conclusions

A novel model-based approach for neuromorphic event camera-based sensors is presented and discussed. The mathematical model of the neuromorphic sensor and its event camera are derived and successfully correlated, relating the amount of slip and pressing force to the event camera's output aperture area. Experimental data from a robotic manipulator-based experiment were used to validate the derived mathematical model. The first-order dynamical behavior of the sensor was verified and validated over its operational dynamical range. However, nonlinear behavior is also observed during the sensor's saturation range, which is primarily attributed to the design limitations of the markers and the use of highly flexible, nonlinear Ninjaflex material. Additionally, the experimental work verified that the proposed sensor could effectively report dynamical information about the three preassigned robotic manipulator actions: approaching, pressing, and slipping.

Temporal and spatial–temporal model-based analysis features were successfully used to train LSTM neural networks to classify the first three preassigned robotic actions: approaching, pressing, and slipping. Once an incipient slip is identified, and with the aid of the newly proposed weighted centroid feature reported by the camera events located in the sensor’s ROI, it is possible to estimate the object’s slip rate. Using the centroid changing rate, one can easily calculate the centroid displacement, which is a direct measure of the object’s slip displacement. This instantaneous slip estimation can be used as a feedback signal for various gripper controller types, enabling precise control commands to secure grip stability and maintain the object’s safety.

Finally, for both the LSTM temporal and the spatial–temporal model based, once the pressing action is identified, a FL-based classifier was successfully used to qualitatively categorize the pressing force extent as “very small,” “small,” “small medium,” “medium,” “small big,” and “big”. Estimating the dynamic information of this pressing force is essential for advanced robotic gripper control strategies.

Acknowledgments. The authors acknowledge and appreciate the assistance provided by the Advanced Research and Innovation Center team at Khalifa University, especially Hussain Sajwani, for his support in setting up the experimental platform.

Author contributions. Murana Awad carried out Conceptualization, Formal analysis, Software Coding/visualization, and Writing – original draft, Review/editing. Musa Abdalla performed Conceptualization, Formal analysis, Mathematical model derivation, Writing – original draft, and Review/editing. Mohammad I. Awad performed Project administration and review. Yahya Zweiri performed Technical advising and writing, including Review/editing. Kinda Khalaf, Project manager, Funding acquisition, and Writing – review/editing.

Financial support. This work was supported by Khalifa University of Science and Technology under Award No. CIRA-2021-106 and by the Advanced Research and Innovation Center (ARIC), jointly funded by Khalifa University of Science and Technology and STRATA Manufacturing PJSC (a Mubadala company).

Competing interests. The authors declare that they have no conflicts of interest.

Ethical approval. Not applicable.

Use of AI. None.

References

- [1] Z. Kappassov, J.-A. Corrales and V. Perdereau, “Tactile sensing in dexterous robot hands – review,” *Robot. Auton. Syst.* **74**, 195–220 (2015).
- [2] Q. Li, O. Kroemer, Z. Su, F. F. Veiga, M. Kaboli and H. Ritter, “A review of tactile information: Perception and action through touch,” *IEEE Trans. Robot.* **36**(6), 1619–1634 (2020).
- [3] H. Yousef, M. Boukallel and K. Althoefer, “Tactile sensing for dexterous in-hand manipulation in robotics – a review,” *Sens. Actuators A Phys.* **167**(2), 171–187 (2011).
- [4] T. Li, Y. Yan, C. Yu, J. An, Y. Wang and G. Chen, “A comprehensive review of robot intelligent grasping based on tactile perception,” *Robot. Comput.-Integr. Manuf.* **90**, 102792 (2024).
- [5] R. A. Romeo and L. Zollo, “Methods and sensors for slip detection in robotics: A survey,” *IEEE Access* **8**, 73027–73050 (2020).
- [6] P. Xiong, Y. Huang, Y. Yin, Y. Zhang and A. Song, “A novel tactile sensor with multimodal vision and tactile units for multifunctional robot interaction,” *Robotica* **42**(5), 1420–1435 (2024).
- [7] S. Zhang, J. Shan, B. Fang and F. Sun, “Soft robotic finger embedded with visual sensor for bending perception,” *Robotica* **39**(3), 378–390 (2021).
- [8] M. Ohka and K. Sagawa, “Stochastic resonance aided tactile sensing,” *Robotica* **27**(4), 633–639 (2009).
- [9] S. Zhang, Z. Chen, Y. Guo, W. Wang, J. Sun, H. Xu, F. Sun, Y. Yang and B. Fang, “Hardware technology of vision-based tactile sensor: A review,” *IEEE Sens. J.* **22**(22), 21410–21427 (2022).
- [10] N. Funk, H. Eberhardt, E. Chalvatzaki, G. Calandra and R. Peters, “EVETAC: An event-based optical tactile sensor for robotic manipulation,” *IEEE Trans. Robot.* (2024).
- [11] M. Li, T. Liu and J. Yao, “Marker displacement method used in vision-based tactile sensors – From 2-D to 3-D: A review,” *IEEE Sens. J.* **28**(8), 8042–8059 (2023).
- [12] W. Yuan, S. Dong and E. Adelson, “Gelsight: High-resolution robot tactile sensors for estimating geometry and force,” *Sensors* **17**(12), 2762 (2017).

- [13] C. Chorley, C. Melhuish, T. Pipe and J. Rossiter, “Development of a Tactile Sensor Based on Biologically Inspired Edge Encoding,” **In: *Proceedings of IEEE International Conference on Robotics and Automation*** (2009) pp. 1–6.
- [14] B. Ward-Cherrier, N. Pestell and N. Lepora, NeuroTac: A neuromorphic optical tactile sensor applied to texture recognition (2020). *arXiv preprint* [arXiv:2003.00467](https://arxiv.org/abs/2003.00467).
- [15] H. Sajwani, A. Ayyad, Y. Alkendi, M. Halwani, Y. Abdulrahman, A. Abusafieh and Y. Zweiri, “TactiGraph: An asynchronous graph neural network for contact angle prediction using neuromorphic vision-based tactile sensing,” *Sensors* **23**(14), 2023 (2023).
- [16] K. Shimonomura, “Event-Based Tactile Image Sensor for Detecting Spatio-Temporal Fast Phenomena in Contacts,” **In: *Proceeding of the IEEE World Haptics Conference (WHC)*** (2019) pp. 343–348 .
- [17] A. Yamaguchi and G. Atkeson, “Tactile behaviors with the vision-based tactile sensor FingerVision,” *Int. J. Hum. Robot.* **16**(3), 1940002 (2019).
- [18] R. Sui, L. Zhang, Q. Huang, T. Li and Y. Jiang, “A novel incipient slip degree evaluation method and its application in adaptive control of grasping force,” *IEEE Trans. Autom. Sci. Eng.* **21**(3), 2454–2468 (2024).
- [19] K. L. Johnson, *Contact Mechanics*, 9th edn. (Cambridge University Press, Cambridge, 2003).
- [20] V. Müller, T.-L. L. Nguyen and E. Veit, “Sensor Design and Model-Based Tactile Feature Recognition.” **In: *Proceedings of the IEEE Sensor*** (2017) pp. 1–4.
- [21] S. U. Hameed, M. Rajkumar, G. Dongming, Z. Yahya and S. Lakmal, “On the design and development of vision-based tactile sensors,” *J. Intell. Robot. Syst.* **102**, 1–27 (2021).
- [22] S. Dong, M. Danielczuk and A. Rodriguez, “Maintaining Grasps within Slipping Bound by Monitoring Incipient Slip,” **In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*** (2019) pp. 3818–3824 .
- [23] D. J. Inman, *Vibration Engineering* (Prentice Hall, Englewood Cliffs, 1996).
- [24] R. D. Howe and Z. Liu, “Beyond Coulomb: Stochastic friction models for practical grasping and manipulation,” *IEEE Robot. Autom. Lett.* **8**(8), 5140–5147 (2023).
- [25] F. Marín, F. Alhama and J. Moreno, “Modelling of stick-slip behavior with different hypotheses on friction forces,” *Int. J. Eng. Sci.* **60**, 13–24 (2012).
- [26] L. A. Levin, P. L. Kaufman and M. E. Hartnett, *Adler’s Physiology of the Eye* (Elsevier, USA, 2024).
- [27] D. del Río and C. Betancourt, “Study of the Human Eye Working Principle: An Impressive High Angular Resolution System with Simple Array Detectors.” **In: *Proceedings of the IEEE Sensor Array Multichannel Signal Processing Workshop (SAM)*** (2006) pp. 1–5.
- [28] K. Shimonomura, “Tactile image sensors employing camera: A review,” *Sensors* **19**(18), 3933 (2019).
- [29] G. Atkeson and A. Yamaguchi, “Recent progress in tactile sensing and sensors for robotic manipulation: Can we turn tactile sensing into vision?,” *Adv. Robot.* **33**(14), 661–673 (2019).
- [30] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis and D. Scaramuzza, “Event-based vision: A survey,” *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(1), 154–180 (2020).
- [31] O. Faris, M. I. Awad, M. Awad, Y. Zweiri and K. Khalaf, A novel bioinspired neuromorphic vision-based tactile sensor for fast tactile perception (2024). *arXiv preprint* [arXiv:2403.10120](https://arxiv.org/abs/2403.10120).
- [32] iniVation, “Understanding the performance of neuromorphic event-based vision sensors,” *White Paper* (2024) [Online]. Available: https://docs.inivation.com/_static/white_papers/White-Paper-May-2020.pdf.
- [33] H. Morrow and R. Kokernak, *Statics and Strength of Materials* (Prentice Hall (Pearson), USA, 2010).
- [34] G. V. Houdt, C. Montero and G. Nápoles, “A review on the long short-term memory model,” *Artif. Intell. Rev.* **53**, 5929–5955 (2020).
- [35] M. O. Abdalla, T. Musa and A. Jafar, “Fuzzy logic control of an electrical traction elevator,” *Jordan J. Mech. Ind. Eng.* **5**(1), 97–106 (2011).
- [36] M. Abdalla, T. Musa and A. Jafar, “Optimal fuzzy controller: Rule base optimized generation,” *Control Eng. Appl. Inform.* **20**(2), 76–86 (2018).
- [37] N. F. Lepora, “Soft biomimetic optical tactile sensing with the TacTip: A review,” *IEEE Sens. J.* **21**(19), 21131–21143 (2021).