
Article:
Xu, X., Saal, H.P. and Ward-Cherrier, B. (2026) TouchReal: an online neuromorphic tactile framework for biomimetic afferent spike generation based on TouchSim. Neuromorphic Computing and Engineering. ISSN: 2634-4386

https://doi.org/10.1088/2634-4386/ae98b0

Neuromorphic Computing and Engineering

PURPOSE-LED PUBLISHING™

TouchReal: an online neuromorphic tactile framework for biomimetic afferent spike generation based on TouchSim

To cite this article before publication: Xingchen Xu *et al* 2026 *Neuromorph. Comput. Eng.* in press https://doi.org/10.1088/2634-4386/ae98b0

Journal Name

Crossmark

RECEIVED
dd Month yyyy

REVISED
dd Month yyyy

PAPER

TouchReal: An Online Neuromorphic Tactile Framework for Biomimetic Afferent Spike Generation based on TouchSim

Xingchen Xu[1], Hannes P. Saal[2] and Benjamin Ward-Cherrier[1,*]

[1]School of Engineering Mathematics and Technology, University of Bristol, Bristol, UK
[2]School of Psychology, University of Sheffield, Sheffield, UK
[*]Author to whom any correspondence should be addressed.

E-mail: b.ward-cherrier@bristol.ac.uk

Keywords: Neuromorphic robotics, force and tactile sensing, biomimetic systems

Abstract

Robotic tactile sensors usually provide continuous, sensor-specific signals such as force, pressure, or vibration. Although such signals are useful for many perception tasks, they are not directly compatible with afferent-level neural representations used in biological touch or spike-based processing systems. Here, we present TouchReal, an online neuromorphic tactile framework that converts signals from a low-cost multimodal tactile sensor into biomimetic SA1-, RA-, and PC-like spike trains inspired by the neural simulator package TouchSim. The hardware combines a capacitive SingleTact force sensor for static contact information with a PVDF piezoelectric film for dynamic vibration-sensitive responses. Raw sensor signals are processed through a pre-processing pipeline and stateful leaky integrate-and-fire (LIF) afferent models. The generated spike trains reproduce several classical temporal, frequency-dependent, and receptive-field-like response properties of cutaneous afferents at an average absolute latency of 130.02 ms with a 32 ms update interval. Functional verification further shows that SA1 and RA spikes support force regression, while PC spikes enable texture classification with around 90% accuracy using temporal encoding features. These results suggest that TouchReal provides an interpretable afferent-level tactile representation for biomimetic robotic sensing, opening up possibilities for neuroprosthetic feedback and robotic models of human perception.

1 INTRODUCTION

Tactile sensing is critical for humans and robots to perceive and manipulate their environment effectively [1] [2]. Researchers have investigated tactile sensing for decades, leading to a wide range of technologies for transducing mechanical interactions into electrical signals [3] including piezoresistive [4], piezoelectric [5] [6], capacitive [7], magnetic [8], and vision-based sensors [9] [10] [11]. These sensors typically provide continuous-valued measurements such as normal force, pressure, shear, deformation, or vibration, which are then processed for task-level perception or control.

In contrast, human tactile perception relies on spike-based, afferent-specific neural representations [12]. Different classes of tactile afferents, such as SA1, RA, and PC[1] fibers, encode complementary aspects of contact interactions and are largely decoupled in both temporal scale and information type [13]. RA afferents respond strongly to dynamic contact events, motion across the skin, and low-frequency vibration (5-50 Hz). SA1 afferents are associated with sustained pressure (0.4-15 Hz), local form, curvature, and fine spatial detail. PC afferents have large receptive fields and high sensitivity to high-frequency vibration and rapid mechanical transients (50-700 Hz), meaning that it can detect sudden changes in pressure and contact related information [2] [14] [15]. These afferent classes do not simply provide redundant copies of the same tactile signal; rather, they decompose contact interactions into complementary channels with different temporal bandwidths, spatial scales, and functional roles. Furthermore, tactile perception is not determined by a single afferent response, but by activity patterns across populations of mechanoreceptive

[1]These afferents are also frequently referred to as SA1/Merkel, RA/Meissner (or FA1), and PC/Pacinian (or FA2) in the literature. To maintain consistency and align with the nomenclature adopted by the TouchSim simulator, we will exclusively use the terms SA1, RA, and PC throughout this paper.

fibers [16] [17] [18]. For example, perception of intensity is mostly based on the activity of all mechanoreceptive afferents [19]. This biological organization suggests that an afferent-level representation could provide a useful interface between artificial tactile hardware and spike-based, human-inspired tactile computation.

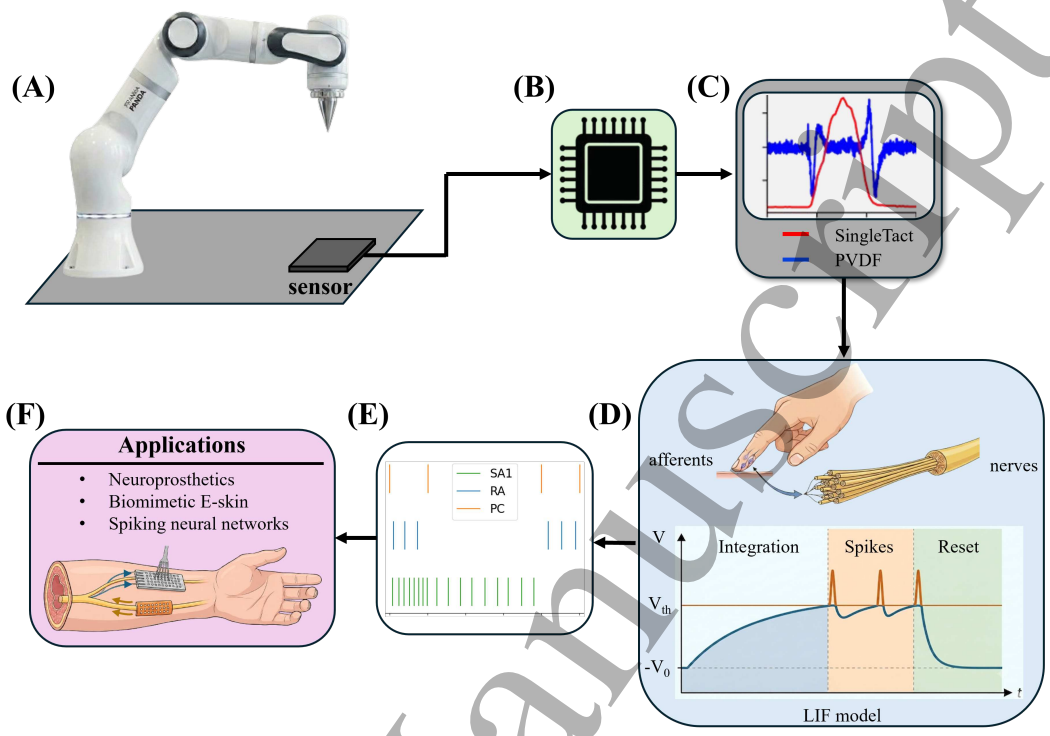

Figure 1. Pipeline overview. (A) The robot arm (Franka Emika Panda) delivers controlled tactile stimuli to the multi-modal e-skin sensor. (B) Raw sensor signals are conditioned by preprocessing circuits and streamed via a microcontroller. (C) An online preprocessing module converts the signals into static and dynamic force components. (D) All force channels drive a stateful TouchSim-inspired LIF afferent model to generate human-afferent-level spike trains. (E) SA1, RA, and PC spike trains are reconstructed online. (F) The spike-based representation enables downstream applications.

Motivated by this idea, several studies have explored the conversion of analog tactile signals into spike-based or event-based tactile representations. Computational receptor models can simulate the responses of tactile afferents under controlled stimuli [20]. Neuromorphic tactile sensors such as NeuroTac generate event-based tactile outputs from optical sensing hardware [21] [22]. More recently, an artificial fingertip [23] composed of FSRs realizes the spike patterns of SA1 afferent with Izhikevich model [24]. At the algorithm level, recent reviews have covered a wide range of methods to generate spikes [25]. Zhao et al. [26] optimize spike encoding and SNN computation for tactile perception. General SNN studies also show the potential of spike-based learning [27].

These studies demonstrate the promise of spike-based tactile processing, but several limitations remain. Some approaches operate only on simulated or highly controlled signals, some produce generic spike streams without explicit correspondence to mechanoreceptive afferent classes, and others focus primarily on task-level inference rather than preserving the functional separation between static pressure, low-frequency dynamic contact, and high-frequency vibration, which are central to biological tactile coding [1]. Consequently, how to convert real, noisy, and sensor-specific tactile measurements into afferent-level spike trains with interpretable SA1-, RA-, and PC-like response patterns remains an open problem.

TouchSim provides a physiologically grounded model for simulating peripheral tactile responses with millisecond precision and afferent-specific structure [28], and has been used previously to model afferent responses. Delhaye et al. [29] use TouchSim to investigate the encoding of geometric features in simulated populations of peripheral tactile afferents; Shao et al. [30] use measured skin displacement signals to drive TouchSim and study dynamic tactile information in the hand, demonstrating the value of combining physical measurements with afferent simulation; Mazzoni et al. [31] use TouchSim to transform scanned texture profiles into mechanoreceptor population responses. These studies show the value of TouchSim-like afferent representations for analyzing tactile coding, perception, and prosthetic feedback. However, TouchSim is primarily designed for

offline simulation with idealized skin mechanics and prescribed stimuli. Directly applying it to robotic tactile sensing is non-trivial, because real sensors introduce noise, compliance, hysteresis, mechanical coupling, sampling-rate constraints, and latency. Therefore, a gap remains between biophysical afferent simulation and online tactile sensing with physical robotic hardware.

In this paper, we propose an online tactile sensing system (Fig. 1) that generates human-afferent-level spike trains compatible with TouchSim-inspired models. We collect ramp and sine indentation data and optimize an afferent model to reconstruct SA1, RA and PC spike responses. We also compare the generated responses both with TouchSim and with known response properties of human mechanoreceptors, and we evaluate whether the generated spikes preserve task-relevant information through two verification tasks.

2 METHODOLOGY

2.1 Tactile Sensing Hardware

This study aims to transform raw multimodal tactile signals into biologically plausible spike trains resembling the responses of three major classes of cutaneous afferents: slowly adapting type I (SA1), rapidly adapting type I (RA), and rapidly adapting type II afferents (PC). These afferents are associated with different frequency ranges where sensitivity is highest, approximately 0.4–15 Hz for SA1, 5–50 Hz for RA, and 50–700 Hz for PC afferents [12]. To approximate these complementary response characteristics, the tactile sensing platform combines a capacitive SingleTact force sensor for static and slowly varying contact force measurements with a PVDF piezoelectric film sensor for dynamic vibration-sensitive measurements.

The SingleTact sensor is a circular capacitive force sensor with an 8 mm diameter. It is operated over a force range of 0–10 N with a resolution of 0.01 N and sampled at 100 Hz. The PVDF sensor (LDT0-028K) consists of a 28 μm-thick piezoelectric PVDF film with screen-printed silver electrodes laminated onto a 0.125 mm polyester substrate. Since the PVDF signal contains higher-frequency dynamic information relevant to RA and PC responses, it is sampled at 2 kHz.

The tactile skin assembly is shown in Fig. 2. The SingleTact sensor and PVDF film are embedded between two silicone layers with thicknesses of 0.5 mm and 1 mm, respectively. This layered structure provides compliant mechanical coupling between the external stimulus and the sensing elements. The assembled skin is mounted on a rigid 3D-printed platform.

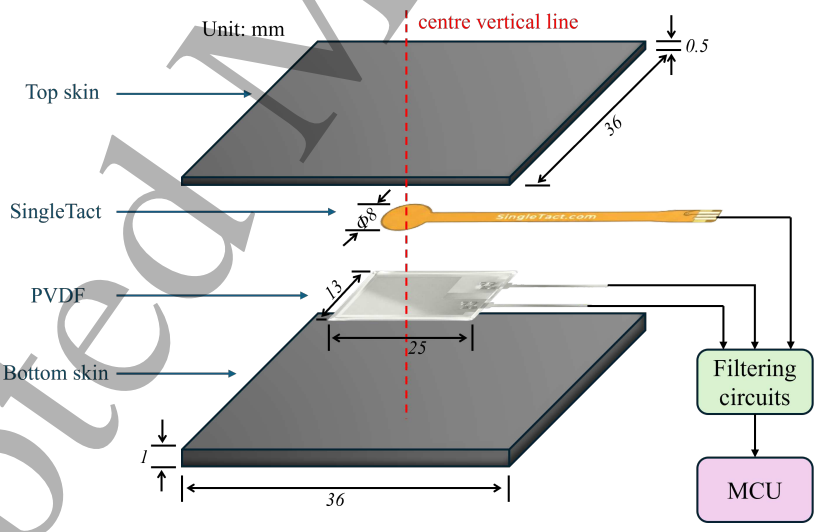

Figure 2. E-skin assembly and signal acquisition pipeline. The SingleTact sensor (yellow) and PVDF film sensor (light gray) are embedded between two silicone layers (dark gray). The sensor outputs are conditioned by filtering circuits and digitized by a microcontroller.

2.2 Data Collection

Ramp and sinusoidal indentation stimuli are collected to train and evaluate the proposed tactile-to-spike encoding framework. Ramp stimuli are used to evaluate static and transient force encoding and are delivered using a rigid probe with a 1 mm diameter mounted on a robotic manipulator (Franka Emika Panda). Sinusoidal stimuli are used to characterize frequency-dependent responses of dynamic afferent channels and are generated using a sound exciter (Dayton Audio DAEX32EP-4 Thruster Exciter, 40 W, 4 Ω). Sinusoidal waveforms with predefined frequencies and amplitudes are generated digitally and amplified using a TPA3116

mono-channel power amplifier before being applied to the exciter, enabling control of the frequency of vibration of the stimulation.

For ramp-shaped stimuli, 448 parameter combinations are collected with 6 trials each by varying indentation velocity, release velocity, hold time, and indentation depth:

- Indentation velocity (V_{up}): 8, 40, 65, and 95 mm/s.
- Release velocity (V_{down}): 8, 40, 65, and 95 mm/s.
- Hold time (h_t): 0, 0.6, 1.2, and 1.8 s.
- Indentation depth (d): 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0 mm.

For sinusoidal stimuli, 30 trials are gathered at each of the following frequencies:

- Frequency: 10, 100, 200, 300 and 400 Hz.

The test dataset uses the same parameter ranges as the training dataset, but consists of independent repeated trials that are not used during parameter optimization. For each physical stimulus condition, corresponding ground-truth afferent spike trains are generated using TouchSim with matched stimulus parameters and trial counts. For sinusoidal stimulation, the indentation is not fixed across frequencies. This is because the effective displacement transmitted to the tactile skin depends strongly on the frequency-dependent coupling between the exciter, probe, silicone layer, and sensor assembly. In particular, higher-frequency vibrations can suffer larger mechanical energy loss, making precise control of physical indentation amplitude difficult. Therefore, before collecting the sinusoidal dataset, the drive amplitude at each frequency is empirically adjusted so that the PVDF response amplitude is approximately matched across frequencies. The frequency-response analysis should therefore be interpreted as the afferent-like firing response under matched sensor-level vibration amplitude, rather than under identical commanded displacement amplitude. The corresponding TouchSim reference stimuli are generated using an indentation depth of 0.1 mm. This TouchSim amplitude is used as a nominal reference level for generating frequency-dependent afferent responses, rather than as a direct estimate of the physical displacement delivered to the sensor at each frequency.

2.3 Signal Preprocessing and Online Implementation

In the original TouchSim formulation [28], tactile input is represented by static, first-order dynamic, and second-order dynamic components. In this work, the representation is adapted to the available multimodal sensing hardware. The preprocessing pipeline converts the raw SingleTact and PVDF recordings into four channels used by the afferent transduction model:

- *ST*: the filtered SingleTact force.
- *d(ST)*: the temporal derivative of the filtered SingleTact force.
- *PVDF*: the reconstructed PVDF signal.
- *d(PVDF)*: the temporal derivative of the reconstructed PVDF signal.

These 4 channels allow the model to represent both slow skin deformation and rapid vibration-induced dynamics, and the filtering is a key component of the pipeline to enable real-world sensor data to be processed.

A causal, chunk-wise pipeline is implemented. At each online update, a newly acquired segment of raw data is passed through stateful preprocessing modules. The sliding window has a length of 128 ms and updates every 32 ms. The 128 ms causal window determines the temporal context required before a stable PVDF-derived output becomes available, whereas the 32 ms update interval determines how often new spike outputs are produced once streaming operation has begun. Thus, the system produces outputs at 31.25 Hz. The raw SingleTact data are processed by a low-pass filter and a causal $\alpha - \beta$ estimator to yield a smoothed ST channel and a temporal derivative *d(ST)* channel.

The PVDF signal contains stimulus-dependent spectral content over a broad frequency range with ramp stimuli mainly producing low-frequency components with broadband noise, whereas high-frequency sinusoidal stimuli produce narrow-band spectral peaks (Fig. 3 A, B). To extract stimulus-relevant dynamic information, the raw PVDF data are processed using a causal frame-based short-time Fourier transform (STFT) and an overlap-add (OLA) front end. Incoming

PVDF samples are accumulated in an input buffer and analyzed once a complete frame is available. It first goes to a low-pass filter and a comb-notch filter to address 50 Hz mains and its harmonics. Each frame is Hann-windowed, then transformed into the frequency domain, and passed through a frequency-dependent manually calibrated gate function (Fig. 3 C). In each STFT frame, spectral components exceeding the calibrated threshold are retained, while weaker components are attenuated. The filtered spectrum is then transformed back to the time domain using the inverse STFT (ISTFT). The reconstructed *PVDF* channel is then differentiated in temporal space to obtain the *d(PVDF)* channel.

Finally, all four channels are resampled to 5 kHz using piecewise cubic Hermite interpolation (PCHIP) to match the temporal resolution used by TouchSim and are passed into a stateful LIF implementation.

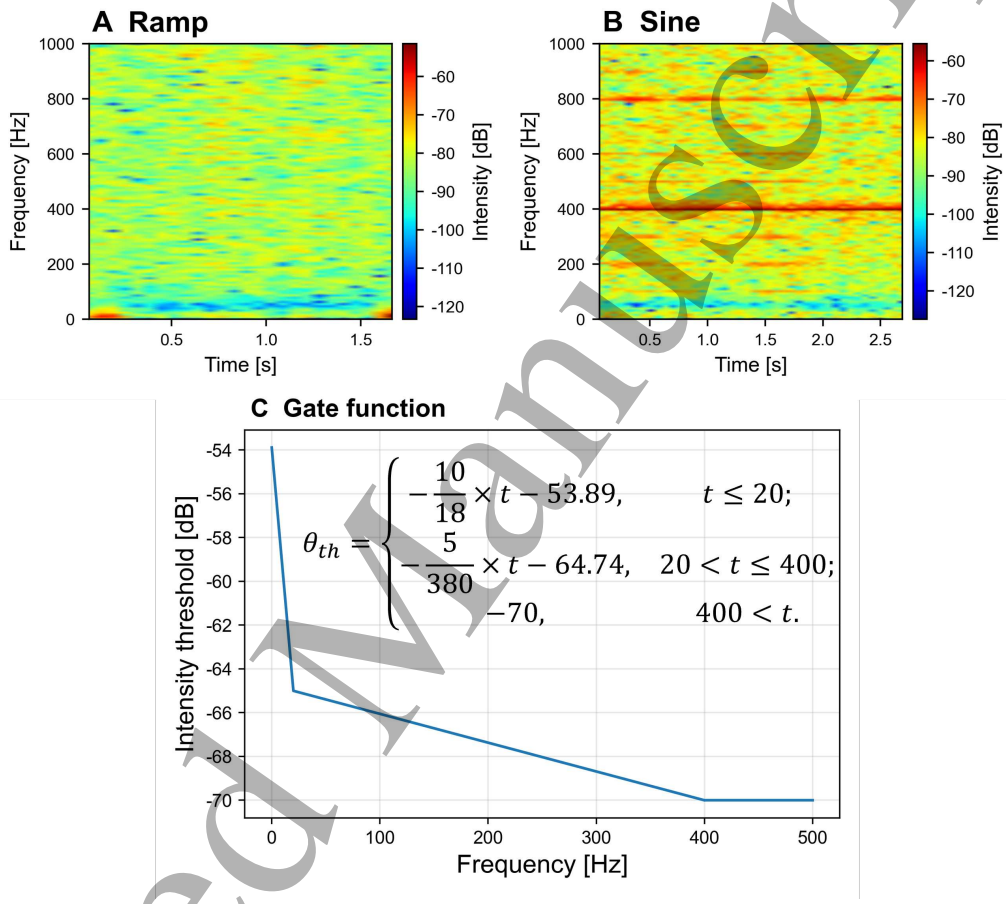

Figure 3. (A) (B) STFT of PVDF outputs of ramp stimulus and a 400 Hz sinusoidal stimulus. The X axis represents time, and the Y axis represents frequency. High-intensity regions correspond to stimulus-related spectral components. (C) Frequency-dependent gate function used for PVDF spectral filtering. In each STFT frame, spectral components stronger than the threshold are retained to preserve stimulus-related dynamics while suppressing noise.

2.4 Parameter Optimization

The afferent transduction stage maps the four previously described preprocessed channels (*ST*, *d(ST)*, *PVDF*, *d(PVDF)*) into SA1-, RA-, and PC-like spike trains. The model (Fig. 4) is adapted from the spiking model of TouchSim. These channels are linearly weighted using afferent-specific parameters and converted into input currents that drive LIF neurons.

For each afferent type, the injected current is computed as a weighted combination of the four input channels providing complementary tactile information. The *ST* channel represents slowly varying contact force and is therefore expected to contribute primarily to sustained SA1-like activity. The *d(ST)* channel emphasizes changes in contact force, including indentation and release transients, and is expected to contribute more strongly to RA-like responses. The *PVDF* channel contains vibration-sensitive dynamic information, while *d(PVDF)* further emphasizes rapid changes and high-frequency transients. These two PVDF-derived channels are therefore expected to contribute most strongly to PC-like responses, while also providing transient information to the RA-like channel. The resulting current is passed through a LIF neuron model with post-spike

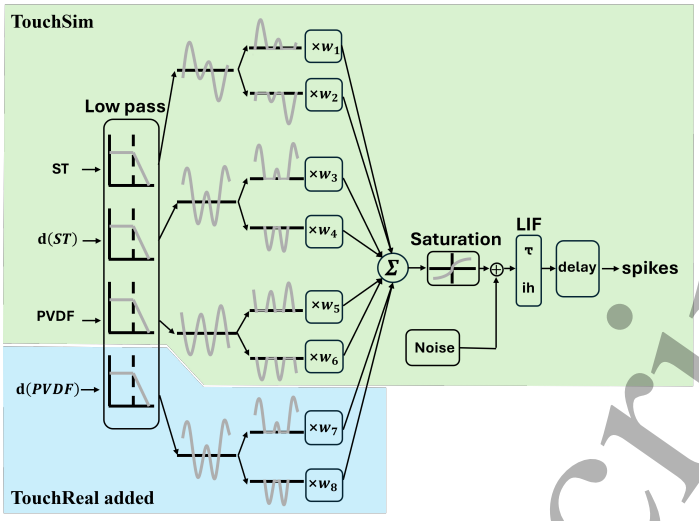

Figure 4. Modified TouchSim spiking model. The four input channels correspond to the SingleTact output, the SingleTact derivative, the reconstructed PVDF signal, and the PVDF derivative. These channels are weighted and combined to drive afferent-specific LIF neurons. The green part replicates the original TouchSim model architecture and the blue part is a novel and necessary contribution of the TouchReal model for real-world applications.

current dynamics inherited from TouchSim. This structure allows SA1-like responses to be dominated by slowly varying force, RA-like responses to emphasize transient force changes, and PC-like responses to depend more strongly on high-frequency PVDF-derived components.

To adapt the model parameters (w_1-w_8, τ, ih, delay) to the proposed sensor configuration, they are optimized by minimizing an objective function. The objective function combines spike timing similarity, spike count consistency, empty-train avoidance, and worst-condition regularization:

$$L = D_{basic} + L_{rel} + L_{empty} \ + L_{MGD} \tag{1}$$

This objective encourages the model to match spike timing, preserve spike count consistency, avoid degenerate silent responses, and maintain acceptable performance across difficult stimulus conditions. A detailed breakdown of this equation is provided in the next sections.

2.4.1 Basic Spike Train Distance To quantify spike train similarity during optimization, we use a normalized kernel-based distance derived from the Van Rossum distance [32]. The original Van Rossum distance can vary over several orders of magnitude depending on spike count and stimulus duration, which can make parameter optimization unstable. Therefore, we use a bounded similarity score.

Let the ground-truth spike train be $S_{gnd} = \{t_i\}$ and the predicted spike train be $S_t = \{u_j\}$. The auto-kernels are defined as:

$$\begin{aligned} A(S_{gnd}) &= \sum_{i=1}^{|S_{gnd}|} \sum_{j=1}^{|S_{gnd}|} exp(-\frac{|t_i - t_j|}{\tau}) \\ A(S_t) &= \sum_{i=1}^{|S_t|} \sum_{j=1}^{|S_t|} exp(-\frac{|u_i - u_j|}{\tau}) \end{aligned} \tag{2}$$

The cross-kernel is:

$$C(S_{gnd}, S_t) = \sum_{i=1}^{|S_{gnd}|} \sum_{j=1}^{|S_t|} exp(-\frac{|t_i - u_j|}{\tau}) \tag{3}$$

The normalized similarity is then:

$$sim(S_{gnd}, S_t) = \frac{C(S_{gnd}, S_t)}{\sqrt{A(S_{gnd})A(S_t)} + \epsilon}, \ \epsilon = 10^{-12} \tag{4}$$

The kernel distance is defined as:

$$D_{kernel}(S_{gnd}, S_t) = 1 - sim(S_{gnd}, S_t) \tag{5}$$

This distance is bounded between 0 and 1, where 0 indicates identical spike trains and larger values indicate weaker similarity.

For each stimulus condition r, distances are computed across all pairs of predicted and ground-truth trials. The basic distance term is:

$$D_{basic} = \frac{1}{n_r n_t n_{gnd}} \sum_{r=1}^{n_r} \sum_{i=1}^{n_t} \sum_{j=1}^{n_{gnd}} D^r_{kernel}(S^i_{gnd}, S^j_t) \tag{6}$$

where n_r is the number of stimulus conditions, n_t is the number of predicted trials per condition, and n_{gnd} is the number of ground-truth trials per condition.

2.4.2 Unbalanced and Empty Spike Train Penalties Kernel similarity alone may not sufficiently constrain firing rate, especially when the optimizer converges to sparse or near-empty spike trains. To encourage spike count consistency, we introduce a relative spike-count penalty. Let N^i_T denote the mean spike count of the ground-truth trials for condition i, and N^i_S denote the mean spike count of the predicted trials. The relative spike-count loss is:

$$L^i_{relative} = \frac{|N^i_T - N^i_S|}{N^i_T + \epsilon}. \tag{7}$$

An additional empty-train penalty is applied when the ground-truth response contains spikes but the predicted response is nearly silent because of the degeneration in convergence:

$$L^i_{\text{empty}} = \begin{cases} \lambda_e L^i_{relative}, & N^i_T \geq thres_1 \,\&\, N^i_S \leq thres_2; \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

In this study, $thres_1 = 1$, $thres_2 = 0.8$, and $\lambda_e = 5$. The average relative count loss and empty-train loss are:

$$L_{rel} = \frac{1}{n_r} \sum_{i=1}^{n_r} L^i_{relative}, \quad L_{empty} = \frac{1}{n_r} \sum_{i=1}^{n_r} L^i_{\text{empty}} \tag{9}$$

2.4.3 Maximum Group Distance Penalty During optimization, the model may fit easy stimulus conditions while neglecting difficult conditions. To reduce this imbalance, we introduce a maximum group distance penalty. For each optimization step, the m stimulus conditions with the largest D^r_{kernel} values are selected, and their mean distance is added to the objective:

$$L_{MGD} = \frac{1}{m} \sum_{r \in \{MGD\}}^{m} D^r_{kernel} \tag{10}$$

where m denotes the set of the m worst-performing stimulus conditions. In this study, $m = 3$.

To assess the functional role of the learned inputs, the additive pre-saturation current is decomposed into the four weighted input components. For afferent type a, the effective contribution of channel k is calculated as

$$C_{a,k} = \frac{\sum_r \sum_t |I^{(r)}_{a,k}(t)|}{\sum_{j=1}^{4} \sum_r \sum_t |I^{(r)}_{a,j}(t)|} \times 100\%, \tag{11}$$

where $I^{(r)}_{a,k}(t)$ is the weighted pre-saturation current contributed by channel k at time t in trial r. Absolute magnitudes are used to prevent positive and negative transient currents from canceling. Contributions are calculated before the joint saturation nonlinearity, where the four channel components remain additively separable.

The same objective structure is used for all three afferent classes to maintain a consistent evaluation criterion, but SA1, RA, and PC models are optimized independently using their corresponding training data and separate parameter sets. The fixed values $thres_1 = 1$, $thres_2 = 0.8$, $\lambda_e = 5$, and $m = 3$ are selected empirically to balance the relative contributions of the four objective terms and prevent any single term from dominating the optimization. These values are implementation choices for the present dataset rather than universal physiological constants.

2.5 Evaluation of Afferent Response Features

After parameter optimization, the model is evaluated by comparing sensor-derived spike trains with TouchSim-generated ground-truth spike trains. The evaluation focuses on three aspects: spike-train mapping ability, classical temporal response features, and spatial receptive-field-like response patterns.

2.5.1 Model Comparison For this evaluation, conventional Van Rossum distance [32] is used to quantify the similarity between sensor-derived and TouchSim-generated spike trains on test data. For each afferent type and stimulus condition, we compute three types of pairwise distances: TouchSim vs TouchSim, TouchReal vs TouchReal, and TouchSim vs TouchReal. This comparison separates the intrinsic variability of TouchSim-generated responses, the variability of sensor-derived responses, and the mismatch between the two response sources.

For this evaluation, the Van Rossum time constant is set to $\tau = 0.001\,\mathrm{s}$, corresponding to a 1 ms temporal resolution. This high temporal precision is used to assess spike timing consistency, while firing-rate-based analyses are additionally used for dynamic afferents whose responses may be strongly affected by small phase shifts.

2.5.2 Temporal Response Features Ramp stimuli are used to examine whether the three afferent channels fire at biologically plausible phases of contact. For each ramp condition, spike trains are averaged across repeated trials to obtain time-varying firing trends. SA1-like responses are expected to dominate during indentation and sustained holding phases, RA-like responses during indentation and release transients, and PC-like responses near high-acceleration turning points.

For sinusoidal stimuli, the average firing rates of RA and PC channels are computed across stimulus frequencies using the PVDF-amplitude-matched trials described previously. This analysis evaluates whether the generated dynamic afferent channels preserve frequency-dependent response trends when the sensor-level vibration amplitude is kept approximately comparable across frequencies. Because millisecond-scale Van Rossum distance can over-penalize small phase shifts in dynamic responses, firing-rate curves are used as a complementary metric for frequency-dependent response validation. Two-way ANOVA is then applied to test the effects of stimulus frequency, response source, and their interaction on firing rate.

2.5.3 Receptive-Field-Like Spatial Response Features Spatial response mapping is performed using ramp/sinusoidal-shaped indentation stimuli applied at different radial offsets from the sensor center. Radial offsets range from 0 to 15 mm with a 1 mm interval, and angular positions range from 0° to 330° with a 30° interval. Each spatial condition is repeated five times. For ramp stimuli, the indentation depths are 0.1, 0.3, 0.5, 0.7, 0.9, and 1.1 mm. Indenting speed (10 mm/s) and hold time (1 s) are set to be constants through the experiment.

For all three afferent-like channels, the generated spike trains are summarized as mean spike count per trial at each spatial sampling position. For the SA1 and RA channels, ramp-shaped indentation stimuli are used because these channels are mainly associated with indentation-related and transient contact events. For the PC channel, sinusoidal stimuli are applied at different radial offsets from the sensor center, with frequencies of 35, 90, 165, 265, and 365 Hz. Although PC responses are reported as firing rates in the frequency-response analysis, spike counts are used here to keep the receptive-field-like spatial analysis consistent across afferent types.

The resulting spatial response maps are interpolated to visualize receptive-field-like response structures. Radial profiles are then computed by averaging responses across positions with the same radial offset. In addition, the area of the valid-response region is estimated by thresholding the interpolated maps using an absolute activity threshold (two spikes). To examine whether the estimated area depends strongly on this threshold, an additional threshold-sweep analysis is performed by varying the absolute spike-count threshold (1-20 mean spike counts). This area is used as an auxiliary descriptor of spatial response spread, while radial attenuation profiles are treated as the primary measure of spatial selectivity.

In addition to radial averaging, angular anisotropy is quantified from the sampling locations using the second-order index:

$$AI_2(r,c) = \frac{|\sum_\theta R(r,\theta,c)e^{i2\theta}|}{\sum_\theta R(r,\theta,c)}, \tag{12}$$

where $R(r,\theta,c)$ denotes the mean spike count at angular position θ under condition c. Opposite directions are treated as the same spatial axis. Large values indicate stronger angular

non-uniformity. The AI_2 value is obtained by averaging across valid nonzero radial rings. For the polar plots, responses are normalized within each radial ring, valid rings are combined using mean spike count as weights, and the resulting profiles are averaged across stimulus conditions.

2.6 Verification Tasks

To evaluate whether the generated biomimetic spike trains preserve task-relevant tactile information, two verification tasks are performed: continuous force regression and texture classification (Fig. 5). These tasks are designed to assess whether task-relevant information is retained after conversion into afferent-like spike trains, rather than to demonstrate superior decoding performance over the original sensor signals. The spike representation is the intended output of TouchReal for biologically grounded neuromorphic processing and potential neuroprosthetic applications.

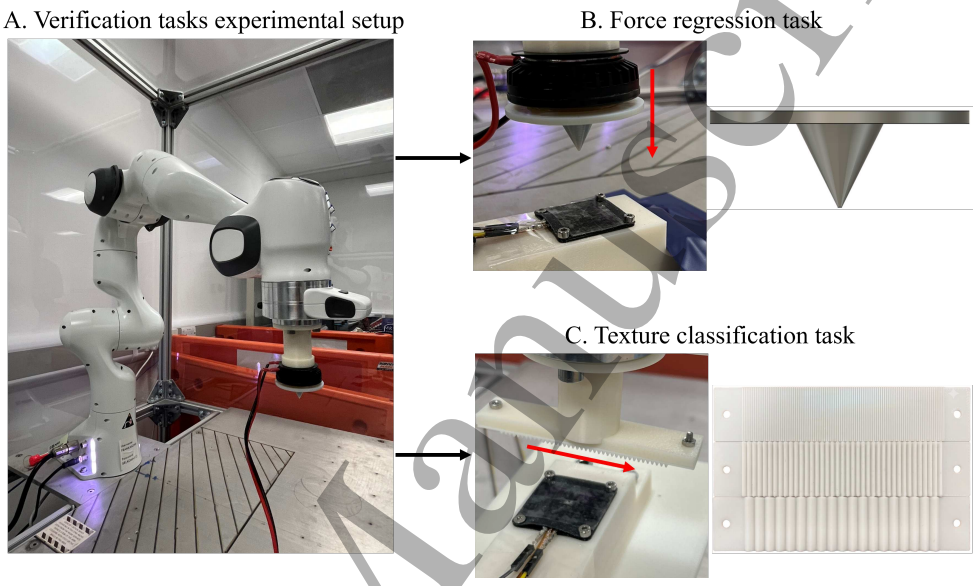

Figure 5. Setup for verification tasks. (A) Robot arm (Franka Emika Panda) with attached stimulus end-effector. (B) Force regression task stimulator mounted on the robot above the sensor. (C) 3D-printed texture classification stimuli with 3 different edge spacings.

2.6.1 Continuous Force Estimation Continuous force estimation is performed using spike trains generated from ramp stimuli. The regression operates online alongside the spike-generation pipeline. Spike times are converted into spike counts using 10 ms bins, and the corresponding ground-truth force label is computed as the mean force within each bin.

At each 32 ms system update, the most recent sliding window of 128 ms is used to provide temporal spike-count features based on online spike generation. A lightweight streaming multilayer perceptron (MLP) is trained to predict force from the spike-count features.

To evaluate the contribution of dynamic afferent information, we compare a model using only SA1 spike trains, only RA spike trains to one using both SA1 and RA spike trains. This comparison tests whether RA-like transient responses improve force reconstruction during dynamic phases such as indentation and release.

Prediction smoothness is also quantified using an oscillation metric based on the second temporal difference of the predicted force:

$$oscillation = \sqrt{\frac{1}{N}\Sigma(\Delta^2\hat{y}_i)^2} \tag{13}$$

where $\Delta^2\hat{y}_i = \hat{y}_{i+1} - 2\hat{y}_i + \hat{y}_{i-1}$. The lower the value, the smoother the curve.

To evaluate the latency-performance trade-off, force regression is additionally repeated across causal processing windows from 4 to 256 ms using the same data partition, spike features, and regression procedure.

2.6.2 Texture Classification Texture classification is performed using comb-shaped artificial textures with three tooth intervals: 0.5 mm, 2 mm, and 4 mm. The indentation depth and sliding

speed are kept constant across textures. The generated PC-like spike trains are used as the primary source of vibration-related texture information.

To reduce the possibility that the classifier learns fixed trial timing rather than texture-dependent spike structure, spike trains are temporally aligned using the first PC spike as the new temporal origin. Spike trains are then binned using 10 ms bins.

Two categories of spike-based features are extracted:

- Rate encoding data: total spike count, average firing rate, inter-spike interval (ISI) features, frequency-domain features.
- Temporal encoding data: Gaussian-smoothed PSTH features used to reduce sensitivity to spike-level jitter while preserving stable temporal response patterns.

Feature ablation experiments are performed using rate encoding data only, temporal encoding data only, and both features. To classify textures in an online manner, sliding window is used with different window lengths (100, 200, 300, 400, 500, 750 ms). The model generates one prediction once the window buffer is filled. A support vector classifier (SVC) is used for texture classification.

The effect of the upstream causal processing window is additionally evaluated using windows from 4 to 256 ms, while the downstream texture-classification inference window is fixed at 500 ms.

3 RESULTS

3.1 Inspection and Spike-Train Comparison of Generated Responses

We first inspect whether the proposed TouchReal model reproduces afferent-specific temporal response patterns under ramp-shaped indentation stimuli. Fig. 6 shows a representative comparison between the ideal TouchSim reference trajectory and the corresponding robot-executed stimulus recorded by the tactile skin. The measured force trajectory follows the ramp-hold-release profile, but exhibits a temporal delay and smoothing effect. This delay is caused by the compliant silicone structure and mechanical damping of the tactile skin, and represents a systematic difference between idealized TouchSim input and physical stimulation. In addition, the robot cannot instantaneously reach and maintain the commanded velocity over short strokes, which further contributes to the temporal shift and smoothing of the measured force trajectory.

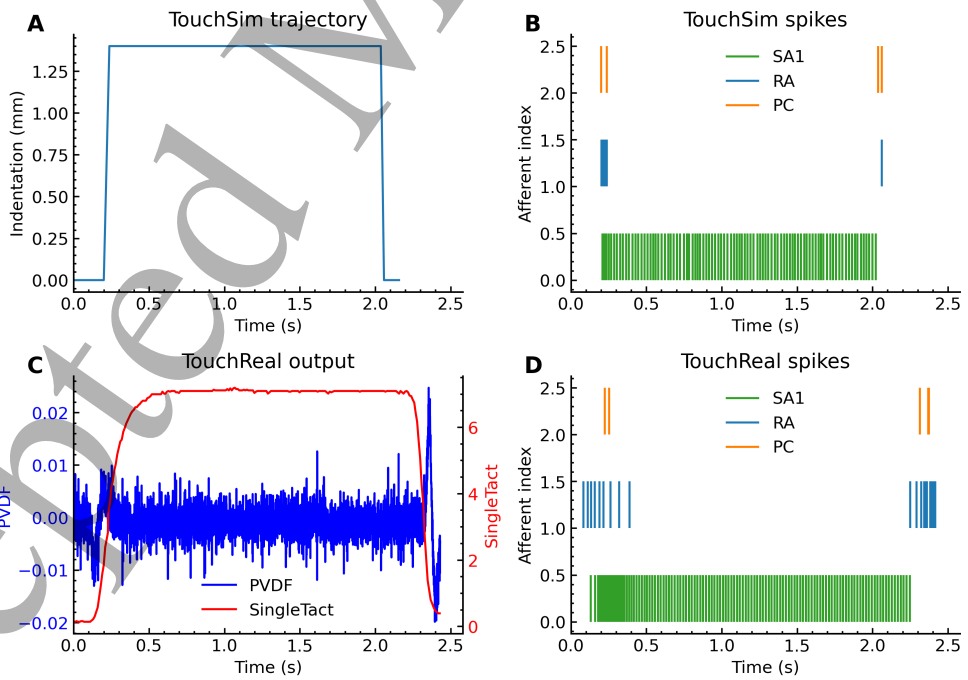

Figure 6. Representative ramp stimulus response. (A) (B) show the TouchSim trajectory and generated spikes. (C) (D) show the corresponding TouchReal output and spikes.

Despite this mechanical delay, the generated spike trains preserve the expected afferent-specific firing phases. SA1-like responses are mainly active during indentation and holding phases, reflecting sensitivity to sustained contact force. RA-like responses are concentrated around indentation and release transients, reflecting sensitivity to temporal changes in contact force.

PC-like responses are primarily observed near rapid transitions, where acceleration and high-frequency components are strongest.

To summarize response trends across repeated trials, spike trains with the same ramp stimulus parameters are averaged over time (Fig. 7). The averaged firing patterns further confirm the functional separation between the three afferent channels. SA1 firing increases during contact formation and remains elevated during the hold phase. RA firing shows transient peaks during indentation and release, while remaining low during the stable holding phase. PC firing is strongest near turning points and rapidly decays outside these high-dynamic events. These temporal response patterns are consistent with known mechanoreceptor response characteristics [33].

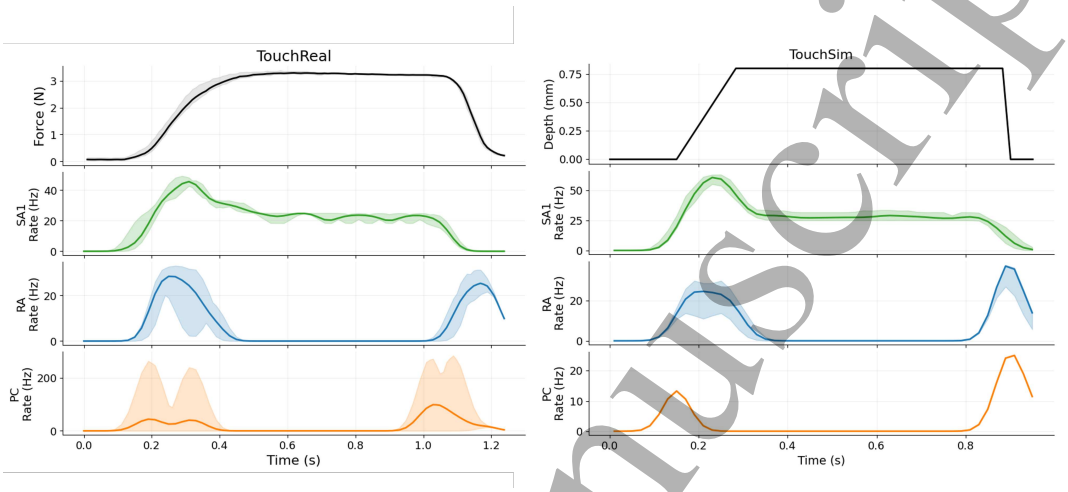

Figure 7. Mean ramp force/indentation and mean afferent firing rates over time of ramp-shaped stimuli for TouchReal (left column) and TouchSim (right column). The top curve shows the measured contact force and indentation depth. The lower panels show averaged SA1, RA, and PC firing rates. Shaded regions indicate the min–max range across repeated trials.

After this qualitative inspection, we quantify spike-train similarity between TouchReal and TouchSim using the conventional Van Rossum distance. For each afferent type, three pairwise comparisons are computed: TouchSim vs TouchSim, TouchReal vs TouchReal, and TouchSim vs TouchReal responses (Fig. 8). This design separates the intrinsic variability of TouchSim reference responses, the variability of sensor-derived responses, and the mismatch between the two sources.

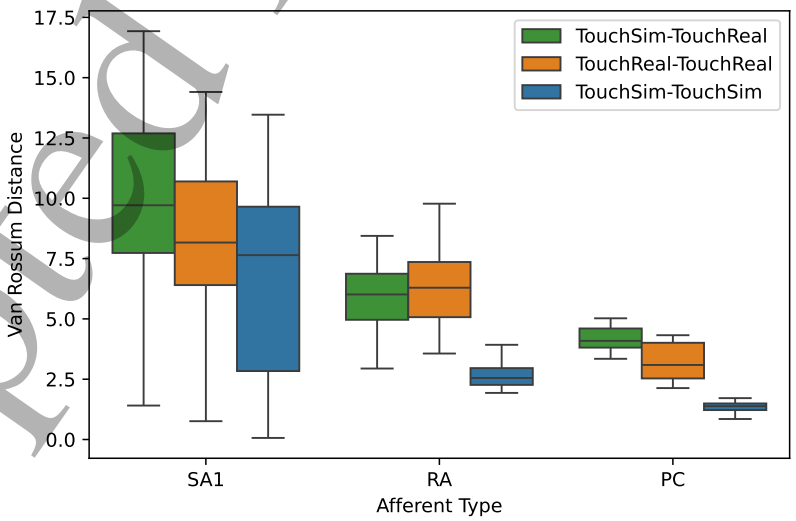

Figure 8. Spike-train mapping evaluation using VR distance. For each afferent type, distances are computed between TouchSim and TouchReal spike trains, between repeated TouchReal spike trains, and between repeated TouchSim spike trains. Boxes indicate the interquartile range (Q1-Q3), center lines indicate medians, and whiskers extend to the most extreme data points within 1.5×IQR.

For SA1 and RA afferents, the inter-model distances (TouchSim–TouchReal) are closer to the intra-model distances (TouchReal–TouchReal and TouchSim–TouchSim) than for PC afferents, indicating that the proposed pipeline reproduces the main temporal response structure of slowly

Table 1. Effective contribution of each input channel

Afferent	*ST*	*d(ST)*	*PVDF*	*d(PVDF)*
SA1	66.79%	27.37%	5.84%	0.00%
RA	0.00%	36.22%	1.26%	62.52%
PC	0.00%	27.19%	44.05%	28.76%

varying and transient tactile responses. The SA1 channel shows larger absolute distances than RA, which is expected because SA1 responses contain longer spike trains during sustained contact; small deviations in sustained firing therefore accumulate over time. Nevertheless, the overlap between the inter-model and intra-model distances suggests that the SA1-like output remains within a comparable response range.

The PC channel shows a larger gap between the TouchSim–TouchReal distance and the two intra-model distances. This indicates that PC-like spike timing is more difficult to reproduce precisely from real sensor data. The mismatch is mainly attributed to the sensitivity of PC responses to high-frequency vibration, acceleration transients, and small temporal alignment errors between simulated TouchSim stimuli and physical contacts for TouchReal. Therefore, for PC responses, spike count and frequency-dependent firing trends are further analyzed as complementary metrics.

To examine whether the learned channel weights are consistent with the intended functional separation, we quantify the effective contribution of each input channel to the pre-saturation injected current. Because *ST*, *d(ST)*, *PVDF*, and *d(PVDF)* have different scales and units, raw weight magnitudes are not compared directly. Instead, the time-integrated absolute magnitude of each weighted channel is calculated and normalized within each afferent type across the test set. As shown in Table 1, the SA1-like current is dominated by *ST*, which accounts for 66.79% of the total effective input, while *d(ST)* contributes a smaller transient component. The RA-like model receives no contribution from *ST*, and 98.74% of its current arises from the two derivative channels. For the PC-like model, the *PVDF* and *d(PVDF)* channels jointly account for 72.81% of the input current, with the remaining contribution arising from *d(ST)*. These results indicate a dominant, although not strictly orthogonal, separation between sustained-force, transient-contact, and vibration-related inputs.

3.2 Temporal Response Features

Because RA and PC afferents are strongly associated with dynamic tactile information, we further evaluate their firing-rate responses under sinusoidal stimulation. Fig. 9 compares TouchSim-generated and sensor-derived firing rates across stimulus frequencies while Tables 2 and 3 quantify the effects of frequency, model source, and their interaction using two-way ANOVA.

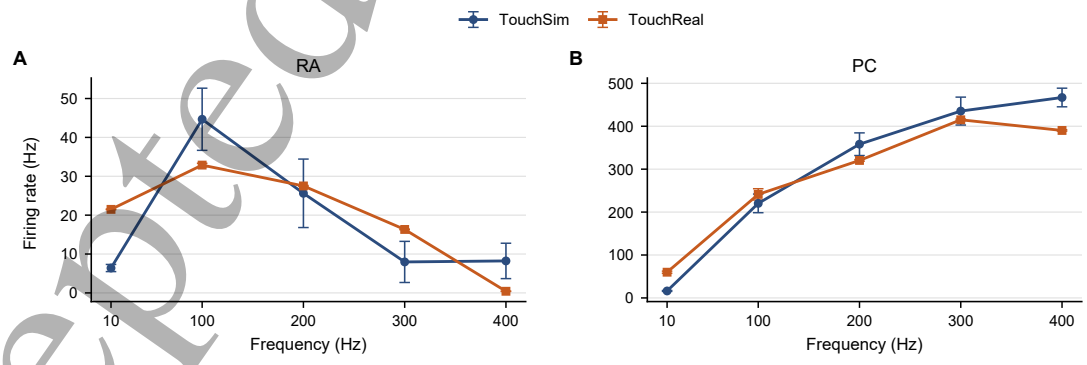

Figure 9. Firing rate versus sinusoidal stimulus frequency under approximately matched PVDF response amplitude. The blue curve represents the TouchSim ground truth, and the orange curve represents the sensor-derived response generated by the proposed TouchReal model. Markers represent trial-averaged firing rates; error bars denote 95% confidence intervals of the mean across trials. (A) RA firing rate's frequency dependency. (B) PC firing rate's frequency dependency.

For RA afferents, both TouchSim and sensor-derived responses show a band-pass-like frequency dependence (Fig. 9 A). Among the tested frequencies, the strongest response occurs at 100 Hz, followed by a decrease at higher frequencies. This trend is supported by the ANOVA, which shows a significant main effect of frequency (Table 2). The main effect of model source is not significant, indicating that the overall firing-rate level of the TouchReal RA channel is not systematically

Table 2. Two-way ANOVA test for RA firing rate

	F	*p-value*	*Significance*
Frequency	46.63	1.15e-32	$p < 0.001$
Model	0.33	0.57	n.s.
Interaction	6.18	7.64e-5	$p < 0.001$

Table 3. Two-way ANOVA test for PC firing rate

	F	*p-value*	*Significance*
Frequency	455.21	7.44e-155	$p < 0.001$
Model	3.43	0.06	n.s.
Interaction	7.92	3.60e-6	$p < 0.001$

shifted relative to the TouchSim reference. However, the significant frequency-by-source interaction shows that the agreement is not uniform across frequencies. This is consistent with the deviations in Fig. 9A, particularly around the response peak and at higher frequencies. These differences are likely to reflect a combination of sensor noise, mechanical damping, and preprocessing-induced changes in the dynamic signal.

For PC afferents, firing rate increases strongly with frequency in both TouchSim and TouchReal (Fig. 9 B). The ANOVA confirms a strong main effect of frequency (Table 3), indicating that the PC-like channel preserves the expected frequency sensitivity. The main effect of model source is not significant, suggesting that there is no strong overall firing-rate bias between TouchReal and TouchSim. The significant interaction term indicates that the frequency-response curves are not identical across all frequencies. This is mainly reflected by the reduced TouchReal response at 400 Hz, where the effective vibration transmitted through the exciter–skin coupling may be weakened.

Overall, both RA and PC responses exhibit strong frequency dependence, which supports that the proposed sensor-to-spike conversion preserves the fundamental frequency sensitivity of tactile afferent responses. The nonsignificant source effects further indicate that the TouchReal responses do not show a systematic overall firing-rate bias relative to the TouchSim reference. The interaction terms show that both afferents have significant frequency-dependent mismatch: The remaining mismatch may reflect frequency-dependent mechanical coupling in the stimulation setup. For example, low-frequency deviations may be related to local resonance between the vibrating tip and the sensor assembly, whereas high-frequency deviations may reflect reduced transmitted vibration energy. The error bars are generally larger for the TouchSim reference than for the TouchReal responses, particularly in the RA channel. This likely reflects the fact that the TouchSim reference includes responses from multiple afferent model instances with different parameter sets, whereas the TouchReal output is generated through a fixed sensor pathway and optimized transduction model. Thus, the TouchSim variability reflects both trial-level variability and the diversity of simulated afferent models, while the TouchReal curve mainly reflects repeatability of the physical sensor-derived response.

3.3 Receptive-Field-Like Spatial Response Features

We next examine the receptive field properties of the generated afferent responses. Spatial analysis is performed by applying ramp indentation and sinusoidal vibration stimuli at different radial offsets and angular positions relative to the sensor center. The mean spike count at each sampling location is used to construct spatial response maps.

Importantly, results shown here should not be interpreted as a direct measurement of biological receptive-field size. In physiological studies, receptive fields are usually defined for individual afferent fibers by mapping their sensitivity profiles, response thresholds, or firing responses across controlled skin stimulation locations. For example, Johansson [34] characterized receptive-field extent as a function of indentation amplitude and described different sensitivity profiles for RA/SA1 and PC/SA2 units, while Vega-Bermudez and Johnson [35] mapped single SA1 and RA afferents using a dense probe array over a fixed skin region. In contrast, the present study uses a single artificial sensor and defines a receptive-field-like area from the spatial region in which the generated spike count exceeds a fixed threshold. Therefore, this metric should be interpreted as a hardware-dependent spatial activation descriptor, rather than as an estimate of human receptive-field size. Its value may vary with sensor geometry, silicone thickness, bonding quality, wire placement, and fabrication variability. Their main significance here is the relative ordering across channels and the change in spatial extent with stimulus condition.

RA responses increase with indentation depth but show a broader spatial distribution than SA1 responses (Fig. 10 A, B). Compared with SA1, the RA maps exhibit stronger peripheral responses and a larger effective response region, suggesting higher sensitivity to dynamic mechanical propagation through the silicone skin. This broader spatial structure is consistent with the role of RA afferents in detecting transient mechanical changes rather than only local static indentation.

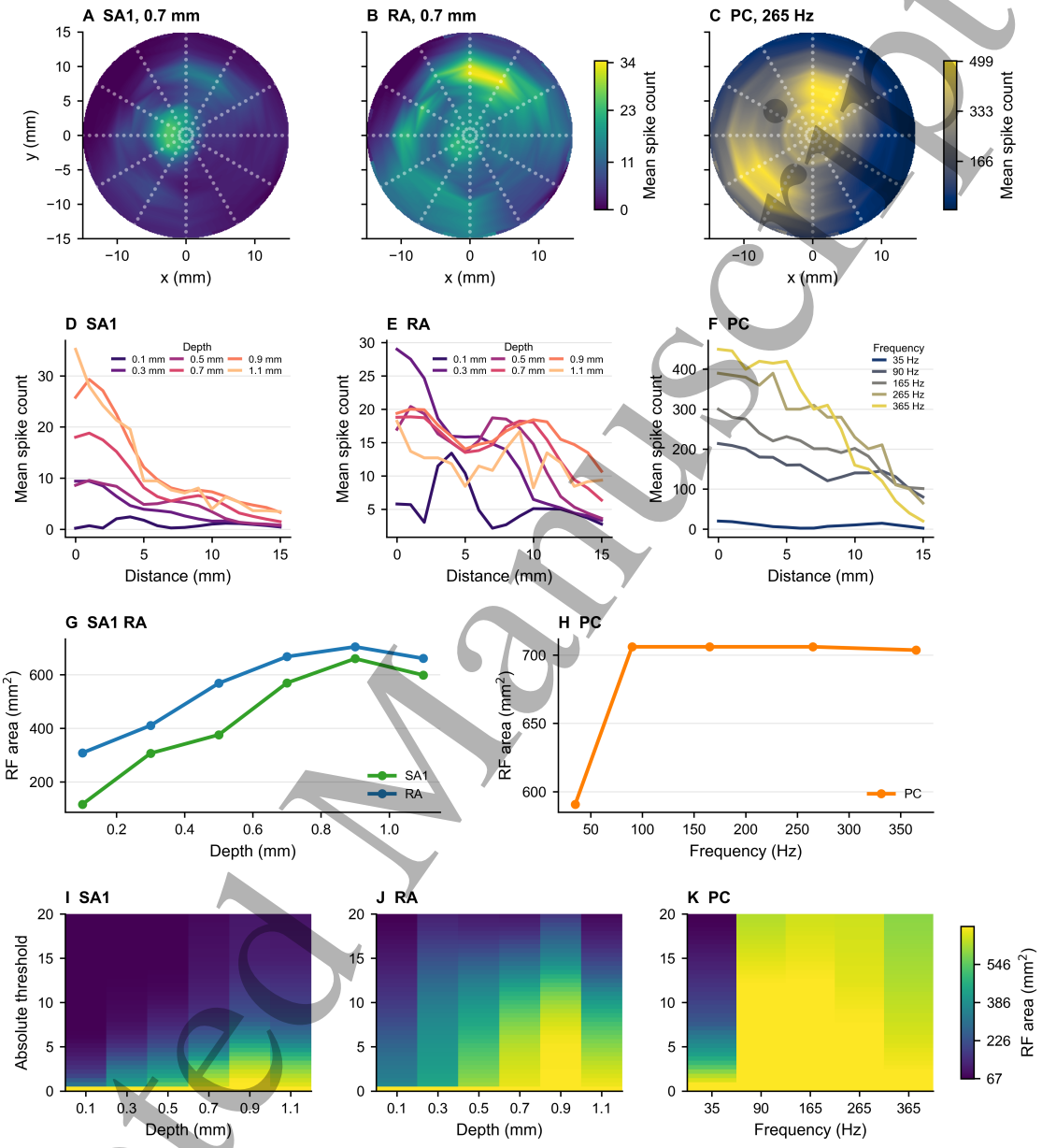

Figure 10. Receptive field features. (A) (B) (C) Examples of spatial response maps. White dots indicate measured sampling locations. Interpolated maps show mean spike counts. SA1 and RA are sampled under 0.7 mm indentation depth. PC is sampled under 265 Hz stimulus. (D) (E) (F) Radial response profiles. Responses are averaged across sampling locations with the same radial offset. All channels are shown as mean spike counts. (G) (H) Effect of indentation depth and vibration frequency on receptive-field-like response area. The area is estimated from interpolated maps using a two-spike threshold. (I) (J) (K) Threshold-sweep heatmap. Receptive-field-like response area changes with different depth and absolute spike count threshold.

For PC responses, the spatial maps are evaluated across sinusoidal stimulus frequencies rather than indentation depths. The PC response increases with frequency and shows a broader vibration-sensitive region at higher frequencies (Fig. 10 C). The maps also exhibit spatial asymmetry, particularly at larger radial offsets. One plausible source of this asymmetry is the mechanical layout of the PVDF sensor. The PVDF film is connected to external wires on one side, and this wired connection may introduce local stiffness, thickness variation, or slight imbalance in the sensor stack. These factors could produce direction-dependent coupling between the probe, silicone layer, and PVDF film, causing vibrations to be transmitted more strongly along some

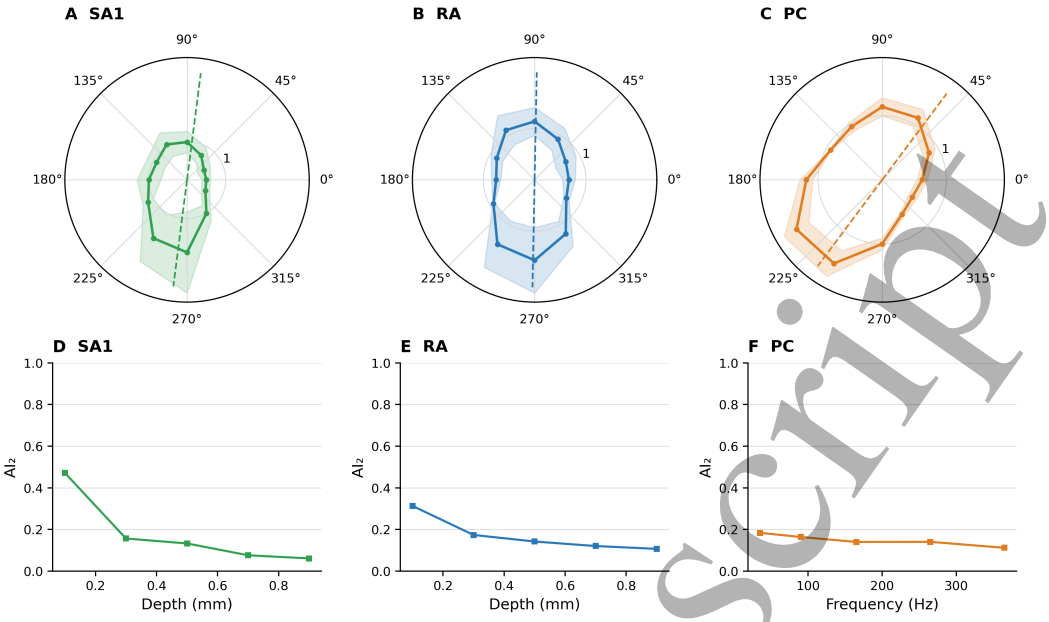

Figure 11. Angular anisotropy of the receptive-field-like responses. (A-C) Normalized angular response profiles for SA1, RA and PC. A value of 1 denotes the angular mean response. Dashed lines indicate the dominant axis. Shaded regions indicate variability across valid radial rings. (D-F) Second-order anisotropy index AI_2 across indentation depth for SA1 and RA and across vibration frequency for PC. Larger values represent stronger angular non-uniformity.

directions than others. Therefore, the observed asymmetry should be interpreted as a feature of the current sensor assembly rather than as a biological PC receptive-field property.

To quantify radial response structure, responses are averaged across angular positions with the same radial offset (Fig. 10, D-F). SA1 responses show a steep radial decay, indicating strong spatial selectivity around the stimulus center. The response amplitude increases with indentation depth, but spatial decay remains relatively sharp. RA responses decay more slowly with radius and show broader response profiles, confirming that the RA-like channel has a larger effective receptive field than SA1. Despite the overall radial decay, RA profiles show secondary peaks at several intermediate radial offsets and depths, which may reflect nonuniform mechanical transmission and local transient dynamics within the silicone-sensor assembly. During off-center indentation, the silicone layer can deform laterally and transmit transient stress to the sensors. Because the sensor stack is not mechanically homogeneous, local differences in bonding, small air gaps, substrate boundaries, and the stiffness of embedded sensor components may cause some off-center stimulation locations, leading to stronger transient signals than neighboring locations. PC responses show high spike counts near the sensor center at higher frequencies, followed by a frequency-dependent decrease with distance. Higher-frequency responses produce stronger central activation but also decay more steeply at larger offsets, whereas lower-frequency responses are weaker but spatially more gradual (Fig. 10 F). This frequency-dependent spatial attenuation is consistent with the idea that tactile vibrations are mechanically filtered as they propagate through compliant tissue or sensor materials. Similar frequency-dependent filtering has been observed in the human hand, where peripheral biomechanics modulate the transmission of vibration signals before they drive PC afferent activity [36]. At larger offsets, PC responses become less regular, likely due to vibration propagation and mechanical asymmetry in the sensor assembly.

Finally, the high-response area is estimated as the interpolated spatial area in which the mean response exceeds two spikes per trial (Fig. 10 G, H). For SA1, the high-response area increases with indentation depth but remains smaller than the RA response area, reflecting stronger localization [15]. RA shows a larger high-response area across most depths, consistent with its broader radial profile. For PC, the high-response area increases with frequency before saturating at approximately 90 Hz. This metric should be interpreted as an auxiliary descriptor rather than a strict physiological receptive field size, because it depends on interpolation, threshold choice, and the peak response value. To check the effect of the spike count threshold, the threshold is varied (Fig. 10, I-K). The threshold-sweep analysis shows that the main trends are preserved across a broad range of thresholds, although the absolute RF-like area depends on the selected threshold.

The angular analysis shows measurable but moderate anisotropy in all three channels (Fig. 11).

For SA1 and RA, AI_2 is highest at the smallest indentation depth and decreases with increasing depth. PC shows relatively stable anisotropy across vibration frequencies, decreasing slightly from approximately 0.18 at 35 Hz to 0.11 at 365 Hz. These results confirm that the spatial responses are not perfectly rotationally symmetric, but are also not dominated by a strong mechanical axis. The moderate PC anisotropy suggests that the rectangular PVDF geometry and asymmetric mounting influence the directional response, while mechanical spreading through the silicone layers and spatial integration across the sensing film reduce strong axis-dependent effects.

3.4 Latency Analysis

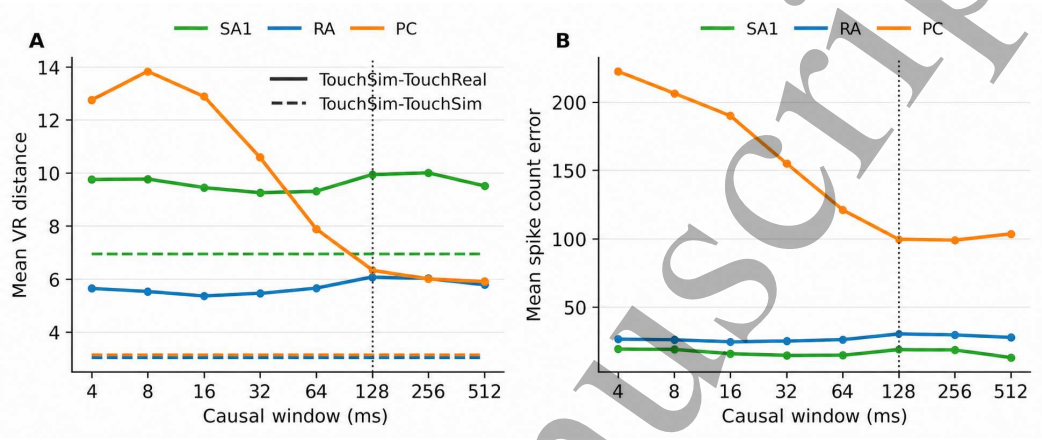

Figure 12. Causal window effect. (A) Solid curves represent latency effect on mean VR distance of different afferents. Dashed lines represent the intra-source TouchSim–TouchSim distance as a baseline. (B) Latency effect on mean spike count error.

Finally, we test different window lengths to examine the effect of latency on model performance. The causal window is varied from 4 to 512 ms, and mean VR distance is computed between TouchReal and TouchSim spike trains (Fig. 12). The overall VR distance decreases rapidly from 4 to 64 ms, indicating that very short STFT windows provide insufficiently stable PVDF reconstruction. Then it decreases with a slower speed, where the actual latency is chosen from. The latency effect is strongly afferent-dependent. PC responses show the largest improvement with increasing latency. It decreases from 12.89 at 16 ms to 6.34 at 128 ms, corresponding to 51% reduction. Further increasing the window to 256 or 512 ms produce only small additional improvements, suggesting that PC channel similarity largely saturates around 128 ms under current configuration. The mean spike count error also shows the plateau around 128 ms. SA1 and RA afferents have rather steady performance across different latencies, suggesting that they have the potential to run in a faster pipeline in the future.

Based on the results, a 128 ms causal processing window is selected as it balances online responsiveness with stable vibration-driven spike-train reconstruction.

As shown in Table 4, with a 128 ms causal window latency, the measured average absolute latency is 130.02 ms, indicating an average processing and scheduling overhead of approximately 2.02 ms beyond the causal window. The 95th percentile latency is 130.67 ms, and the worst observed latency is 136.83 ms. These results indicate that the computational overhead is small compared with the causal processing window which is required for robust PVDF preprocessing. The measured computation time is below the 32 ms update interval, allowing the system to maintain continuous online operation without accumulating delay across successive updates.

The current latency is therefore mainly determined by the signal-processing window required for robust PVDF reconstruction under noisy real-sensor conditions, rather than by the computational cost of the model itself. This suggests that future improvements in sensor signal quality and mechanical integration could reduce the required temporal window and further decrease the end-to-end latency.

3.5 Verification Tasks

3.5.1 Continuous Force Estimation To test whether the generated spike trains preserve force-related information, an online force estimation task is performed using ramp stimulus data. Spike trains are converted into 10 ms spike-count bins, and a causal 128 ms sliding window is used

Table 4. TouchReal Latency

Quantity	Value	Meaning
Update interval	32 ms	Time between successive outputs
Update rate	31.25 Hz	Output frequency
Causal Window	128 ms	Required temporal context
Mean overhead	2.02 ms	Processing and scheduling beyond the window
Mean end-to-end latency	130.02 ms	Total measured latency
95th percentile latency	130.67 ms	Runtime variability
Worst observed latency	136.83 ms	Maximum measured value

Table 5. Force Regression Performance

	RMSE	*MAE*	*Oscillation*	R^2	*Pearson r*
SA1 only	0.6775	0.4591	0.3228	0.9336	0.9685
RA only	2.4203	2.0074	0.5887	0.1525	0.3971
SA1 + RA	0.6115	0.3650	0.2800	0.9459	0.9730

as the input to a lightweight MLP regressor. Three input configurations are compared: SA1 spikes only, RA spikes only and combined SA1+RA spikes.

Table 5 summarizes the regression performance. The SA1-only model achieves an RMSE of 0.6775, an R^2 of 0.9336, and a correlation of 0.9685, showing that SA1-like spikes alone encode most of the sustained force information. Adding RA spikes further improves performance, reducing RMSE from 0.6775 to 0.6115 and increasing R^2 from 0.9336 to 0.9459. The oscillation metric also decreases from 0.3228 to 0.2800, indicating smoother predictions.

Fig. 13 shows representative regression traces. The SA1+RA model more accurately follows rapid changes during indentation and release, while maintaining stable predictions during the holding phase. It also has the lowest RMSE and highest R^2 and Pearson r coefficient, confirming the benefit of combining sustained and transient afferent-like information. This supports the interpretation that SA1 spikes primarily encode sustained force magnitude, whereas RA spikes provide complementary information about dynamic force transitions [33].

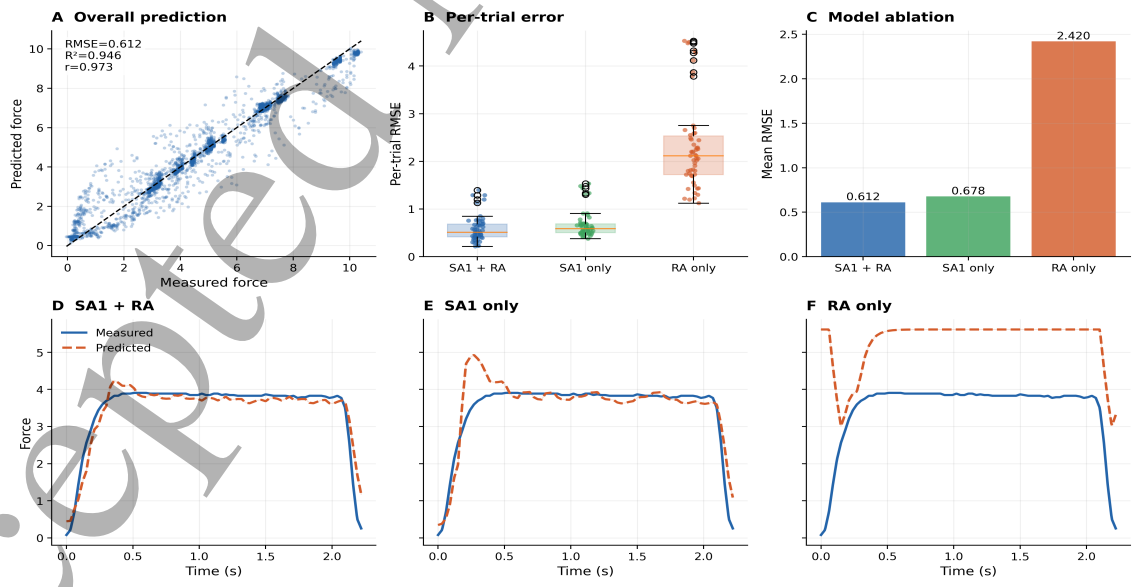

Figure 13. Force regression from spike trains. (A) True and predicted forces of the test set (SA1 + RA) in all sliding windows. The black dashed line indicates the identity line corresponding to perfect prediction. (B) Paired test-trial error distribution. Each dot represents the RMSE of a single test trial. (C) Mean RMSE across test sets. (D) Single trial examples for the SA1+RA, SA1 only and RA only conditions under fixed parameters (depth: 1 mm; indenting speed: 38 mm/s; release speed: 94 mm/s; hold time: 1.8 s).

Force regression performance shows a clear dependence on the available temporal context (Fig. 14). Increasing the causal window from 4 to 128 ms reduces *RMSE* from approximately 2.36

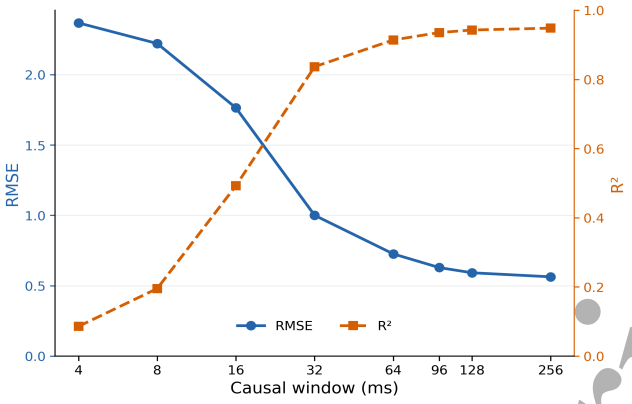

Figure 14. Effect of causal processing window length on force regression performance. The same spike features, data partition, and regression procedure are used for all conditions. Longer windows improved force decoding, with performance approaching a plateau at approximately 96-128 ms.

to 0.61 and increases R^2 from approximately 0.09 to 0.95. Most of the improvement occurs between 16 and 64 ms, while performance approaches a plateau beyond 96-128 ms. These results indicate a clear latency–accuracy trade-off: very short windows reduce absolute latency but provide insufficient temporal information for reliable force reconstruction, whereas windows around 96–128 ms offer substantially more stable decoding performance.

3.5.2 Texture Classification The final verification task evaluates whether PC-like spike trains preserve vibration information. Three comb-shaped artificial textures with different tooth intervals are classified using spike-based features extracted from PC responses. To reduce dependence on absolute trial timing, all spike trains are aligned to the first PC spike before feature extraction.

Predictions are made from PC spike activity within a recent temporal sliding window (Fig. 15). Classification accuracy increases with longer windows, indicating a trade-off between response latency and available temporal evidence. Short windows provide limited information. At 100 ms, the mean temporal-encoding-only accuracy is 57.8%, while rate-encoding-only features achieve 63.9%. This likely reflects the small number of PC spikes available within such a short interval, making the temporal pattern sparse and less reliable. Performance of temporal encoding features improves substantially once the window exceeds 300-500 ms while rate encoding features remain lower and more variable.

Combining both features does not consistently improve performance over temporal encoding alone. This indicates that the temporal encoding representation already captures the dominant discriminative information, while the added rate encoding features provide limited additional benefit and may introduce extra variability.

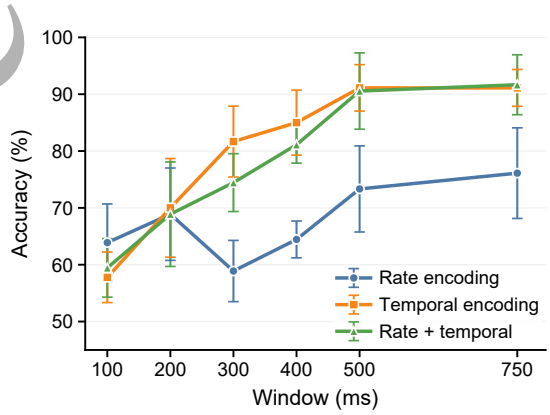

Figure 15. Effect of the inference window length on classification accuracy. Longer windows contain a longer spike-train history and therefore provide more temporal evidence for classification.

The effect of causal processing window length on texture decoding is further evaluated while keeping the downstream inference window fixed at 500 ms (Fig. 16). Classification performance remains close to the 33.3% chance level for the 4 and 8 ms conditions. The highest mean accuracy

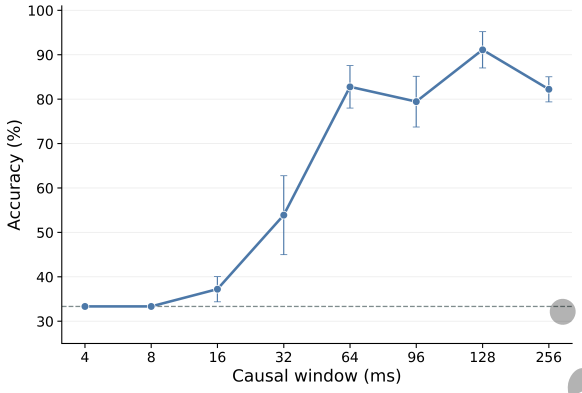

Figure 16. Effect of causal processing window length on texture classification accuracy. Texture features are extracted using a fixed 500 ms inference window from PC-like spike trains generated with causal processing windows ranging from 4 to 256 ms. The dashed line indicates the 33.3% chance level for the three-class classification task.

is obtained at 128 ms, reaching approximately 91%. These results indicate that short causal windows do not provide sufficient temporal context for stable PC-like texture encoding, consistent with the reduced frequency resolution and less stable vibration-feature estimation associated with shorter STFT windows.

4 DISCUSSION

This work presents an online tactile encoding framework that maps tactile sensor signals into biomimetic SA1-, RA-, and PC-like spike trains. Rather than treating tactile sensing as a direct regression or classification problem from continuous sensor signals, the proposed system introduces an intermediate afferent-level representation inspired by TouchSim. The results show that this representation can reproduce several classical temporal and spatial response properties of cutaneous afferents, while preserving task-relevant information for force regression and texture classification. This makes the representation potentially useful for applications that require interpretable, spike-based tactile feedback, such as biomimetic robotic perception and neuroprosthetic interfaces.

4.1 From real sensor signals to afferent-like spike trains

A central challenge in applying TouchSim-like models to robotic tactile sensing is the mismatch between idealized mechanical stimuli and real sensor measurements. TouchSim assumes well-defined indentation trajectories and skin mechanics, whereas the present system receives signals from a compliant silicone-sensor assembly. As shown in the ramp response examples, the measured force follows the intended ramp-hold-release profile but exhibits smoothing and temporal delay, especially during the release phase. This delay is not simply a software artifact; it reflects the mechanical compliance, hysteresis, and damping of the silicone structure and the sensor mounting. Consequently, perfect spike-time alignment with the ideal TouchSim reference is challenging to implement.

Despite this mismatch, the generated spike trains preserve the main afferent-specific temporal structure. These response phases are consistent with the expected functional separation between different afferents. This suggests that the proposed preprocessing and transduction pipeline can recover interpretable afferent-level structure from noisy physical sensor measurements, even when the underlying mechanical trajectory differs from the ideal reference.

The mapping results further show that SA1 and RA channels are more closely matched to the TouchSim reference than the PC channel. This difference is expected. SA1 responses are dominated by slowly varying force and are therefore more robust to small timing errors. RA responses depend on transient force changes, but their relevant temporal scale is still relatively compatible with the SingleTact derivative and low-frequency PVDF components. In contrast, PC-like responses are highly sensitive to vibration amplitude, high-frequency content, and precise stimulus onset. Small variations in exciter-skin coupling, contact timing, and signal thresholding can therefore produce large differences in spike timing, even when the overall firing-rate trend remains similar. Accordingly, the current PC output is interpreted as PC-like based on its combined transient, frequency-dependent, spatial, and texture-related response properties, while exact spike-time correspondence with an ideal TouchSim reference remains limited.

4.2 Optimization under spike-count and spike-timing constraints

The parameter optimization problem is difficult because spike-train similarity depends on both spike timing and spike count. A timing-based distance alone may produce undesirable solutions when spike trains are very sparse, when trial durations differ, or when physical stimulus timing is shifted relative to the TouchSim reference. In particular, channels can converge to near-silent responses if the injected current remains below threshold over many conditions. Such solutions may appear numerically stable but are functionally invalid.

Although the objective contains several terms to avoid these issues, the optimization framework still has limitations. A single loss formulation is used across three channels, while these afferent types encode different temporal features and are affected by different noise sources. Some manual tuning may be necessary, especially when the collected data contain strong noise and when the framework is applied to other hardware platforms. Future work should therefore consider afferent-specific objective functions instead of a universal loss.

4.3 Spatial response structure and receptive-field-like behavior

The spatial scanning experiments show that the generated SA1-, RA-, and PC-like responses exhibit distinct receptive-field-like structures. SA1-like responses are strongly center-dominant and decay rapidly with radial distance, indicating localized spatial selectivity. RA-like responses are broader and show slower radial decay, suggesting stronger sensitivity to dynamic mechanical propagation through the silicone layer. This difference is consistent with the expected contrast between localized slowly adapting responses and more spatially distributed rapidly adapting responses [15].

The PC-like channel shows a broad vibration-sensitive response once the stimulus frequency exceeds the effective activation threshold. This is consistent with the broad receptive-field characteristics typically associated with Pacinian afferents. However, the PC spatial maps also show clear asymmetry, particularly at larger radial offsets. This asymmetry is likely to arise from the physical sensor assembly, including nonuniform silicone bonding, local air gaps, sensor placement, and boundary effects.

The qualitative ordering of receptive-field size observed here ($SA1 < RA < PC$) is consistent with biological tactile afferents. Previous microneurography studies reported SA1 receptive-field areas of approximately 5–9 mm^2 and RA receptive-field areas of approximately 5–22 mm^2 on the human fingertip, depending on stimulus conditions [35] [37]. In contrast, the receptive-field-like areas measured in this study reach several hundred square millimeters (Fig. 10 G, H), with maximum values of approximately 660 mm^2 for SA1-like responses and 710 mm^2 for RA-like responses. This difference is expected because the present measurements reflect the combined effects of mechanical propagation within the silicone layer, sensor transduction, and spike encoding, rather than the receptive field of a single biological afferent. Furthermore, the RF area in this study is defined using supra-threshold spike-count responses, whereas biological receptive fields are typically mapped using near-threshold neural activation. The large spatial scale mainly reflects the dimensions of the present sensor and mechanical spreading through the silicone assembly. Future sensor miniaturization and reduced mechanical coupling length scales may bring the effective areas closer to physiological values. Therefore, the measured RF areas should be interpreted as effective sensor receptive fields rather than direct physiological equivalents.

4.4 Latency relative to biological tactile processing

The current streaming implementation achieved an average absolute latency of 130.02 ms with a 32 ms update interval. This latency is longer than early cortical responses reported in human tactile processing. For example, tactile stimulation can evoke a P50 component at approximately 50 ms, which has been associated with early contralateral somatosensory cortical activity. Hämäläinen et al. reported a first distinct contralateral response at about 50 ms after tactile pulses [38], and Staines et al. reported a vibrotactile P50 latency of 56 ± 7 ms [39]. Human ECoG studies have also shown that S1 high-gamma responses to vibrotactile stimuli can peak within approximately 50–100 ms after stimulus onset [40]. These studies suggest that early human cortical tactile processing occurs on the order of several tens of milliseconds.

Compared with this biological timescale, the present system is slower in terms of absolute latency. However, the measured computational overhead was only approximately 2 ms beyond the designed 128 ms causal window. Thus, the main source of latency is not the LIF transduction model itself, but the temporal window required for robust online preprocessing of noisy PVDF signals. In particular, the STFT–OLA reconstruction and spectral gating require sufficient temporal context to distinguish stimulus-related vibration from broadband noise and power-line

interference. This trade-off is necessary for the current low-cost single-pixel hardware, but it is not a fundamental limitation of the afferent-spike representation.

Although the 128 ms window provides the best overall performance, the verification results at 64 ms remain reasonably strong if not consider the PC-like afferent matching performance, indicating that a shorter operating window may be feasible when reduced latency is prioritized over maximum decoding accuracy. This brings the system closer to biologically relevant tactile-processing timescales, although further validation in closed-loop tasks is required.

Improving the hardware signal quality could therefore substantially reduce the system latency. Better mechanical integration between the silicone layers and sensors, reduced air gaps, improved shielding and grounding, lower-noise amplification, and more stable PVDF mounting would reduce the need for long spectral windows and aggressive denoising. With cleaner dynamic signals, shorter causal windows could be used for PVDF reconstruction, moving the end-to-end latency closer to the early human tactile processing range. This provides an important direction for future work, especially for closed-loop robotic control and neuroprosthetic feedback where low-latency tactile information is critical.

4.5 Functional relevance of the generated spike representation

The two verification tasks indicate that the generated spike trains retain useful tactile information beyond qualitative resemblance to biological afferents. The force regression results are consistent with the known functional roles of tactile afferents in biological touch. Previous studies have shown that SA1 afferents provide a robust representation of sustained contact force, whereas RA afferents respond to force transients during loading and unloading phases [41] [42]. The observation that SA1 spikes alone recover most of the force trajectory, while the addition of RA spikes further improves reconstruction accuracy, agrees with the complementary contributions of slowly adapting and rapidly adapting afferents to force perception and manipulation.

The texture classification results are also consistent with biological evidence for texture coding. Fine textures are known to generate high-frequency skin vibrations that are represented predominantly by PC afferents through precise temporal spike patterns rather than firing rate alone [43] [44]. In this study, temporal encoding features substantially outperformed rate encoding descriptors, suggesting that temporal spike structure carries more texture information than global firing-rate statistics. This finding is consistent with previous observations that temporal coding in PC afferents provides critical information for texture discrimination [45].

Moreover, classification performance increases rapidly as the observation window expands from 100 to 500 ms, reaching approximately 90% accuracy at 500 ms. This timescale is comparable to psychophysical studies showing that humans can reliably identify material and texture properties within a few hundred milliseconds of tactile exploration [33] [46]. The similarity between the temporal requirements of the classifier and human tactile perception suggests that the generated PC-like spike trains preserve relevant texture information rather than merely reproducing afferent-like firing statistics.

These tasks are intended as verification experiments rather than full application benchmarks. Together, these results support the use of afferent-like spikes as an intermediate tactile representation for downstream perception and control, while remaining consistent with established biological mechanisms of force and texture encoding.

4.6 Limitations and future work

Several limitations remain. First, the current hardware is based on a single-pixel tactile sensor. Spatial response maps are obtained by scanning the stimulus across the sensor surface rather than by recording from a dense tactile array. As a result, the measured receptive-field-like patterns reflect both the encoding model and the mechanical properties of the sensor assembly. In contrast, biological tactile perception relies on distributed population activity across many afferents rather than on isolated single-unit responses. Studies have shown that SA1 and RA population responses encode spatial fingertip stimulation patterns, and that ensembles of human tactile afferents can represent complex fingertip events and force-related information [35] [42] [47]. Higher-density tactile arrays would therefore allow simultaneous spatial sensing and more direct population-level afferent encoding, making the artificial representation closer to the distributed coding strategy used in biological touch.

Second, the mechanical coupling between the rigid probe, silicone layers, SingleTact sensor, and PVDF film strongly affects the recorded signals. Small variations in bonding, air gaps, and sensor alignment can change the spatial and frequency response, especially for the PVDF channel.

Improving the mechanical integration of the sensor, for example by better lamination or vacuum-assisted sealing, may reduce nonuniform coupling and improve repeatability.

Third, the PC-like channel remains the most difficult to fit. High-frequency responses are sensitive to vibration amplitude, resonance, and temporal alignment, and the current STFT-based online reconstruction introduces a trade-off between noise suppression, frequency selectivity, and latency. Future work should investigate adaptive spectral processing and afferent-specific optimization objectives to improve PC robustness under real-world stimulation.

Finally, although the present system is evaluated using force regression and texture classification, the broader motivation is to provide a neural-compatible tactile representation for closed-loop robotic control and neuroprosthetic feedback. Extending the framework to tactile arrays, multi-contact manipulation, and closed-loop robot learning would allow the generated afferent-like spikes to be evaluated in more realistic interaction tasks.

5 CONCLUSION

This work presents TouchReal, an online neuromorphic tactile framework that converts multimodal tactile sensor signals into biomimetic SA1-, RA-, and PC-like spike trains. By combining a capacitive SingleTact sensor, a PVDF piezoelectric film, causal signal preprocessing, and a stateful TouchSim-inspired LIF model, the system provides an interpretable afferent-level representation of real tactile interactions. The generated spike trains reproduce key temporal response patterns, frequency-dependent RA and PC firing trends, and receptive-field-like spatial structures. The streaming implementation achieves an average absolute latency of 130.02 ms with a 32 ms update interval, with computation contributing only a small overhead beyond the designed 128 ms causal processing window. Functional evaluations further show that the generated spikes preserve task-relevant information for force regression and texture classification. Although the current system remains limited by single-pixel sensing, mechanical coupling, and the latency required for robust PVDF denoising, the results demonstrate the feasibility of using low-cost tactile hardware to generate biologically interpretable afferent-like spike representations. Future work will focus on reducing latency through improved hardware integration and lower-noise sensing, extending the framework to tactile arrays, and evaluating the spike representation in closed-loop robotic and neuroprosthetic applications.

Funding

This work was supported by the Royal Academy of Engineering Fellowship on "Shared Autonomy Neuroprosthetics: Bridging Artificial and Biological Touch" under Grant RF/202021/20/171 and funded by the China Scholarship Council.

References

[1] R. Romo and E. Salinas, "Touch and go: Decision-making mechanisms in somatosensation," *Annual review of neuroscience*, vol. 24, no. 1, pp. 107–137, 2001.

[2] R. S. Johansson and J. R. Flanagan, "Coding and use of tactile signals from the fingertips in object manipulation tasks," *Nature Reviews Neuroscience*, vol. 10, no. 5, pp. 345–359, 2009.

[3] R. S. Dahiya, G. Metta, M. Valle, et al., "Tactile sensing—from humans to humanoids," *IEEE transactions on robotics*, vol. 26, no. 1, pp. 1–20, 2009.

[4] S. Sundaram, P. Kellnhofer, Y. Li, et al., "Learning the signatures of the human grasp using a scalable tactile glove," *Nature*, vol. 569, no. 7758, pp. 698–702, 2019.

[5] M. Tanaka, K. Miyata, and S. Chonan, "A wearable braille sensor system with a post processing," *IEEE/ASME Transactions On Mechatronics*, vol. 12, no. 4, pp. 430–438, 2007.

[6] H. Hu, Y. Han, A. Song, et al., "A finger-shaped tactile sensor for fabric surfaces evaluation by 2-dimensional active sliding touch," *Sensors*, vol. 14, no. 3, pp. 4899–4913, 2014.

[7] S. Luo, X. Zhou, X. Tang, et al., "Microconformal electrode-dielectric integration for flexible ultrasensitive robotic tactile sensing," *Nano Energy*, vol. 80, p. 105 580, 2021.

[8] Y. Yan, Z. Hu, Y. Shen, et al., "Surface texture recognition by deep learning-enhanced tactile sensing," *Advanced Intelligent Systems*, vol. 4, no. 1, p. 2 100 076, 2022.

[9] B. Ward-Cherrier, N. Pestell, L. Cramphorn, et al., "The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies," *Soft robotics*, vol. 5, no. 2, pp. 216–227, 2018.

[10] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[11] X. Xu, N. F. Lepora, and B. Ward-Cherrier, "A neuromorphic tactile system for reliable braille reading in noisy environments," *IEEE Robotics and Automation Letters*, 2025.

[12] W. H. Talbot, I. Darian-Smith, H. H. Kornhuber, et al., "The sense of flutter-vibration: Comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand.," *Journal of neurophysiology*, vol. 31, no. 2, pp. 301–334, 1968.

[13] K. O. Johnson, S. S. Hsiao, and T. Yoshioka, "Neural coding and the basic law of psychophysics," *The Neuroscientist*, vol. 8, no. 2, pp. 111–121, 2002.

[14] K. O. Johnson, "The roles and functions of cutaneous mechanoreceptors," *Current opinion in neurobiology*, vol. 11, no. 4, pp. 455–461, 2001.

[15] A. B. Vallbo, R. S. Johansson, et al., "Properties of cutaneous mechanoreceptors in the human hand related to touch sensation," *Hum neurobiol*, vol. 3, no. 1, pp. 3–14, 1984.

[16] K. Johnson, "Reconstruction of population response to a vibratory stimulus in quickly adapting mechanoreceptive afferent fiber population innervating glabrous skin of the monkey.," *Journal of Neurophysiology*, vol. 37, no. 1, pp. 48–72, 1974.

[17] H. P. Saal and S. J. Bensmaia, "Touch is a team effort: Interplay of submodalities in cutaneous sensibility," *Trends in neurosciences*, vol. 37, no. 12, pp. 689–697, 2014.

[18] G. Corniani, M. A. Casal, S. Panzeri, et al., "Population coding strategies in human tactile afferents," *PLOS Computational Biology*, vol. 18, no. 12, e1010763, 2022.

[19] M. A. Muniak, S. Ray, S. S. Hsiao, et al., "The neural coding of stimulus intensity: Linking the population response of mechanoreceptive afferents with psychophysical behavior," *Journal of Neuroscience*, vol. 27, no. 43, pp. 11 687–11 699, 2007.

[20] Q. Ouyang, J. Wu, Z. Shao, et al., "A python code for simulating single tactile receptors and the spiking responses of their afferents," *Frontiers in neuroinformatics*, vol. 13, p. 27, 2019.

[21] B. Ward-Cherrier, N. Pestell, and N. F. Lepora, "Neurotac: A neuromorphic optical tactile sensor applied to texture recognition," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 2654–2660.

[22] X. Xu, A. Li, and B. Ward-Cherrier, "Exploratory movement strategies for texture discrimination with a neuromorphic tactile sensor," *arXiv preprint arXiv:2509.14954*, 2025.

[23] M. Statache, M. Filosa, S. Ballanti, et al., "Neuromorphic touch sensors for pleasantness encoding via spike patterns: A bio-inspired artifical fingertip for restoring affective sensations," in *2024 IEEE SENSORS*, IEEE, 2024, pp. 1–4.

[24] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on neural networks*, vol. 14, no. 6, pp. 1569–1572, 2003.

[25] T. Liu, G. Brayshaw, A. Li, et al., "Neuromorphic touch for robotics-a review," *Neuromorphic Computing and Engineering*, 2025.

[26] X. Zhao, X. Wang, J. Wu, et al., "Spiketouch: Optimizing spike neural networks for tactile perception," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 9, no. 4, pp. 1–24, 2025.

[27] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, et al., "Deep learning in spiking neural networks," *Neural networks*, vol. 111, pp. 47–63, 2019.

[28] H. P. Saal, B. P. Delhaye, B. C. Rayhaun, et al., "Simulating tactile signals from the whole hand with millisecond precision," *Proceedings of the National Academy of Sciences*, vol. 114, no. 28, E5693–E5702, 2017.

[29] B. P. Delhaye, X. Xia, and S. J. Bensmaia, "Rapid geometric feature signaling in the simulated spiking activity of a complete population of tactile nerve fibers," *Journal of neurophysiology*, vol. 121, no. 6, pp. 2071–2082, 2019.

[30] Y. Shao, V. Hayward, and Y. Visell, "Compression of dynamic tactile information in the human hand," *Science advances*, vol. 6, no. 16, eaaz1158, 2020.

[31] A. Mazzoni, C. M. Oddo, G. Valle, et al., "Morphological neural computation restores discrimination of naturalistic textures in trans-radial amputees," *Scientific reports*, vol. 10, no. 1, p. 527, 2020.

[32] M. C. van Rossum, "A novel spike distance," *Neural computation*, vol. 13, no. 4, pp. 751–763, 2001.

[33] B. P. Delhaye, K. H. Long, and S. J. Bensmaia, "Neural basis of touch and proprioception in primate cortex," *Comprehensive physiology*, vol. 8, no. 4, pp. 1575–1602, 2018.

[34] R. S. Johansson, "Tactile sensibility in the human hand: Receptive field characteristics of mechanoreceptive units in the glabrous skin area," *The Journal of Physiology*, vol. 281, pp. 101–125, 1978.

[35] F. Vega-Bermudez and K. O. Johnson, "Sa1 and ra receptive fields, response variability, and population responses mapped with a probe array," *Journal of Neurophysiology*, vol. 81, no. 6, pp. 2701–2710, 1999.

[36] N. Tummala, G. Reardon, B. Dandu, et al., "Biomechanical filtering supports efficient tactile encoding in the human hand," *Journal of the Royal Society Interface*, vol. 23, no. 237, 2026.

[37] R. S. Johansson and A. B. Vallbo, "Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin," *The Journal of Physiology*, vol. 286, pp. 283–300, 1979.

[38] H. Hämäläinen, J. Kekoni, M. Sams, et al., "Human somatosensory evoked potentials to mechanical pulses and vibration: Contributions of si and sii somatosensory cortices to p50 and p100 components," *Electroencephalography and clinical neurophysiology*, vol. 75, no. 1-2, pp. 13–21, 1990.

[39] W. R. Staines, C. Popovich, J. K. Legon, et al., "Early modality-specific somatosensory cortical regions are modulated by attended visual stimuli: Interaction of vision, touch and behavioral intent," *Frontiers in psychology*, vol. 5, p. 351, 2014.

[40] S. Ryun, J. S. Kim, H. Lee, et al., "Tactile frequency-specific high-gamma activities in human primary and secondary somatosensory cortices," *Scientific reports*, vol. 7, no. 1, p. 15 442, 2017.

[41] R. S. Johansson and G. Westling, "Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip," *Experimental Brain Research*, vol. 56, pp. 550–564, 1984.

[42] I. Birznieks, P. Jenmalm, A. W. Goodwin, et al., "Encoding of direction of fingertip forces by human tactile afferents," *Journal of Neuroscience*, vol. 21, no. 20, pp. 8222–8237, 2001.

[43] A. I. Weber, H. P. Saal, J. D. Lieber, et al., "Spatial and temporal codes mediate the tactile perception of natural textures," *PNAS*, vol. 110, no. 42, pp. 17 107–17 112, 2013.

[44] Bensmaia, "Biological and bionic hands: Natural neural coding of texture," *Philosophical Transactions of the Royal Society B*, 2014.

[45] Lieber and Bensmaia, "High-dimensional representation of texture in somatosensory cortex," *Nature Communications*, 2020.

[46] Klatzky and Lederman, "Identifying objects from a haptic glance," *Perception & Psychophysics*, 1995.

[47] R. S. Johansson and I. Birznieks, "First spikes in ensembles of human tactile afferents code complex spatial fingertip events," *Nature Neuroscience*, vol. 7, no. 2, pp. 170–177, 2004.