2018 IEEE International Conference on Robotics and Automation (ICRA)
May 21-25, 2018, Brisbane, Australia

Calibration and Analysis of Tactile Sensors as Slip Detectors

Karl Van Wyk, Joe Falco

Abstract— **The existence of tactile afferents sensitive to slip-related mechanical transients in the human hand augments the robustness of grasping through secondary force modulation protocols. Despite this knowledge and the fact that tactile-based slip detection has been researched for decades, *robust* slip detection is still not an out-of-the-box capability for any commercially available tactile sensor. This research seeks to bridge this gap with a comprehensive study addressing several aspects of slip detection. In particular, key developments include a systematic data collection process yielding millions of sensory data points, a spectral analysis of sensory responses providing insight into sensor behavior, and the application of Long Short-Term Memory (LSTM) neural networks to produce robust slip detectors from three commercially available sensors capable of tactile sensing. The sensing mechanics behind these sensors are all fundamentally different and leverage principles in electro-mechanical resistance, optics, and hydro-acoustics. Critically, slip detection performance of the tactile technologies is quantified through a measurement methodology that unveils the effects of data window size, sampling rate, material type, slip speed, and sensor manufacturing variability. Results indicate that the investigated commercial tactile sensors are inherently capable of high-quality slip detection.**

Index Terms— **tactile sensors, slip detection, neural networks, deep learning.**

I. Introduction

Neurophysiological research reveals the existence of four distinct types of tactile afferents in the human hand: fast-adapting type I (FA-I), slow-adapting type I (SA-I), fast-adapting type II (FA-II), and slow-adapting type II (SA-II) [1]. This variety of tactile afferents affords a spectrum of sensitivity to various mechanical stimuli, an arrangement determined critical for proper sensorimotor control of the hand. Among the many functional modalities these afferents impart, one of particular interest is the sensing of high-frequency mechanical transients (5 Hz - 400 Hz) via FA-I and FA-II type afferents. For example, these vibrations can emanate from surfaces in sliding contact that undergo quick transitions among friction states, also known as the ”catch and snap” effect [2]. Naturally, humans possess the ability to detect object slip, and, in fact, humans actively exploit slip detection by applying secondary force modulation efforts in response to slip events with a reflex time of less than 100 msec [3]. In essence, this sensorimotor response acts as a ‘fail-safe’ in the event that primary force modulation protocols do not suffice. Therefore, imbuing robotic hands and grippers with tactile-enabled slip detection behaviors is one logical avenue for elevating the robustness of robotic grasping.

The calibration of tactile sensors for slip detection is not a novel concept; investigative efforts started in the late 1980s and continue to the modern day. To date, researchers have produced slip detectors from wildly different sensors including accelerometers [2], force transducers [4], pressure-sensitive tactile arrays [5], [6], piezoelectric polymer films [7]–[9], elastomer-embedded cameras [10], carbon nanotube-polymer composites [11], pressure transducers [12], capacitance arrays [9], and strain gauges [13]. The majority of slip detector algorithms are at least partially composed of spectral analyses from Fourier and wavelet transforms, to extract relevant features for slip classification [5], [9], [13]. Other approaches apply optical flow algorithms [6], band-pass filters [12], or contact force cone and force measurements to predict slip [14]. Reported slip classification accuracies were above 90 % in [12] and above 85 % in [9].

Despite such a rich variety of sensor designs and algorithmic approaches for slip detection, commercial tactile sensors are still devoid of this capability. The resulting implication is that slip-detecting tactile sensors have not yet approached readiness levels for commercialization. In particular, the calibration of tactile sensors for truly *robust* slip detection across all relevant factors (including slip speed, contacting materials, latent vibrations, and contact force) has not yet been demonstrated. This result is likely due to both methodological and algorithmic inadequacies. Methodologically, slip detection research rarely investigates detection performance across even a subset of the previously listed factors. Algorithmically, subsequent slip detection accuracies are purely insufficient for applied controls. The former issue can be resolved by a thorough design of experiments across all relevant factors in order to systematically hone and test slip detector quality. The latter issue can be addressed with the application of more powerful algorithms capable of analyzing sequences of sensor data for classifying slip events with a high degree of accuracy.

In the era of high performing neural networks in challenging problem domains, strategies for producing highly robust slip detection for tactile sensors should include neural networks and automated data acquisition of large datasets. Coupling deep neural networks and extensive datasets with contemporary computational power have led to significant performance gains in computer vision [15], language modeling [16], language translation [17], artificial gaming agents

K. Van Wyk and J. Falco are with the National Institute of Standards and Technology, Gaithersburg, MD, USA (e-mail: karl.vanwyk@nist.gov; joseph.falco@nist.gov).

Disclaimer: Certain commercial equipment, instruments, or materials are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

[18], bin picking [19], and speech synthesis [20]. Depending on the problem domain, some neural architectures prove advantageous over others. Since slip detection is primarily imparted through temporal signatures of tactile data, the most promising neural architecture for slip detection is *recurrent* neural networks.

By design, recurrent neural networks excel at processing *sequences* of data through the existence of internal neuronal states in the hidden layers that integrate input data over the entire input sequence. Recurrent networks have been notoriously difficult to train in the past due to their exploding or vanishing gradients. However, this issue has been largely resolved by Long Short-Term Memory (LSTM) recurrent networks [21]. LSTM networks circumvent these numerical issues during the training process through their intrinsic dynamics, and excel at finding patterns in data separated by large time intervals - an attractive trait for slip detection where several hundred sensor readings may be necessary for maximizing accuracy.

This research is the first to apply LSTM networks to large datasets gathered over the previously listed factors to create robust slip detectors. The main contributions include: 1) a systematic design of experiments for collecting millions of data points for training slip detectors; 2) an in-depth, comparative spectral analysis of each sensor, yielding insight on the most significant frequencies for slip detection; 3) preprocessing and application of LSTM networks for calibrating a sensor for slip detection; and 4) a measurement methodology that analyzes windows size, sampling rate, material, slip speed, and manufacturing variability effects on slip detection accuracy.

II. Experimentation

A. Hardware

The main hardware components (existing commercial products) included a seven degree-of-freedom (DoF) robotic arm; a fully actuated, sixteen DoF robotic hand (individual torque control and angle sensing per joint); and three tactile sensor technologies. The arm was Cartesian position controlled, and the hand was joint impedance controlled (proportional-derivative controller with gravity model) with a joint stiffness of 3 Nm/rad, joint damping of 0.075 Nm·s/rad, and torque saturation limits of ± 0.42 Nm. Each set of tactile sensors was mounted as the fingertips (i.e., index, middle, little) of a robotic hand attached to the arm as depicted in Fig. 1. Sensor 1 (three-count) was a six-axis, silicon strain gauge-based force-torque transducer with a 3D printed, rubber-coated fingertip and a force sensitivity of 6 mN. Sensor 2 (three-count) was a three-axis force sensor that used optics to measure mechanical deformation of the outer rubber dome with a force sensitivity of 2.5 mN. Sensor 3 (two-count) was a biomimetic sensor with a fluid-filled rubber membrane encapsulating a rigid core that housed multiple electrodes, a pressure transducer, and a thermistor with a force sensitivity of at least 10 mN. Data was collected from all analog signals at 1000 Hz from Sensor 1 and Sensor 2 and from the high-pass filtered pressure transducer signal at 850 Hz for Sensor 3 (the manufacturer-specified signal specifically designed for applications of high-frequency sensory transients).

B. Factors

In order to train a robust slip detector with LSTM networks, a large, highly representative dataset was necessary. Specifically, relevant factors that affect a candidate sensor's response in both slip and non-slip events must be known. These factors include sensor force sensitivity, sampling rate, tribological properties of both sensor surface and extrinsic surface, speed of slip between sensor and extrinsic surface, ambient vibration from the sensor's connecting structure or extrinsic surface, and the sensor's force-loading profile.

C. Data Collection

In order to cover all *controllable* factors, the data collection process was designed to sample both non-slipping and slipping types of data under different scenarios. This process was consistently applied across all three types of tactile sensors (see Fig. 1).

1) Non-Slipping Data: Two scenarios were engaged to collect non-slipping tactile data. The first scenario involved repositioning the arm and hand to 36 different joint configurations in free space (see Fig. 1f) at three different speed settings, yielding average end-of-arm speeds of 25 mm/s, 50 mm/s, and 75 mm/s. Each speed setting was sequentially applied to the same set of 36 joint configurations. Finger joint configurations were generated via the Latin Hypercube sampler with joint velocities pseudo-randomly sampled between 2 deg/s and 10 deg/s. The data generated for this scenario were labeled as non-slipping data since the tactile sensors did not make surface contact. This data aided in training a slip detector that did not issue false-positive slip detection from the idling vibration or motion of the connecting electromechanical structures (i.e., arm and hand).

Another case of non-slipping data included force modulating the sensors at varying frequencies and magnitudes with a prospective surface in order to emulate scenarios in which the robotic hand or gripper readjusted its grasp or performed in-hand manipulation on an object without slipping. The robotic arm positioned the robotic hand within reach of five different material sheets: aluminum, polyvinyl chloride (PVC), neoprene, cardboard, and plywood. The fingers were commanded to initiate contact and palpate the materials at various force levels and frequencies by commanding the fingers into the working surface at 100 different inwardly-curled joint angle configurations and speeds sampled from the Latin Hypercube sampler (see Fig. 1(a-e)). Each joint angle configuration was issued as a change in joint angle per joint across all three contacting fingers ranging from 0 deg to 70 deg (per joint). Joint speeds ranged from 0.125 deg/s to 25 deg/s.

2) Slipping Data: In order to capture the tactile sensory response to slipping events, variations in slip speed, material, and contacting force must be represented. Accordingly, the robotic arm was commanded to drag the hand with tactile-surface contact across various material types 32 times at

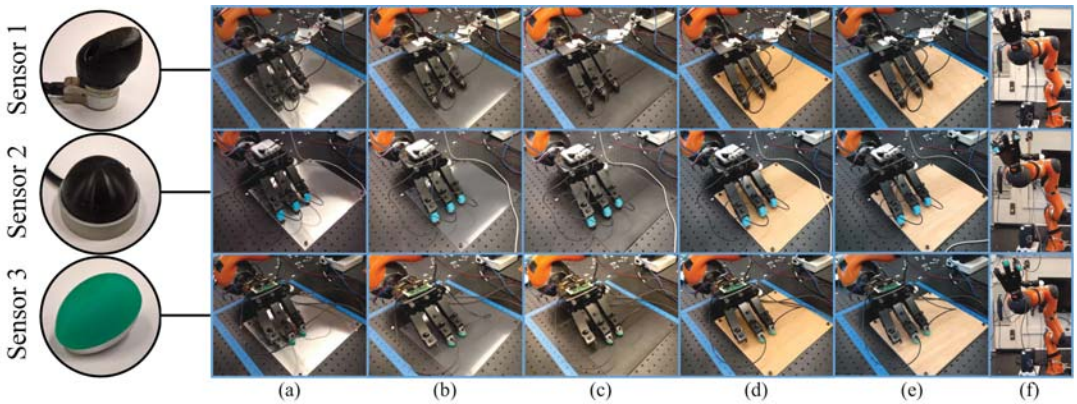

Fig. 1: Gathering tactile data on (a) aluminum, (b) PVC, (c) neoprene, (d) cardboard, (e) plywood, and in (f) free-space.

four speed settings, yielding average end-of-arm speeds of 5 mm/s, 25 mm/s, 50 mm/s, and 75 mm/s (see Fig 1(a-e)). Although only four arm speed settings were used, controller design requires smooth motion trajectories, which generated a continuous spectrum of slip speeds up to the maximum speed setting. Joint angles to the hand were held constant per set of tactile sensors that yielded appreciable contact with the materials. Deviations in hand joint angles were not issued in order to prevent accidental fingertip-surface lift and to reduce wear on the sensor surfaces from excessive contact forces. However, contact force was inherently modified by the natural oscillation of the moving arm along its planned motion path.

D. Data Balancing

After acquisition, the data was balanced and split into training and testing datasets. To prevent a classifier from exhibiting biases towards more frequently represented patterns in the data, non-slipping and slipping data, and sources thereof, were equally represented. Overall, a representative dataset consisted of a 50-50 % split of non-slipping and slipping data. With five materials and four speed settings, each material yielded 20 % of the slipping data, and each slip speed setting yielded 25 % of each material. Moreover, pushing data and free-space data constituted 50 % of non-slipping data, each. Each material yielded 20 % of the pushing data. Each speed setting yielded 33 % of free-space data. To enforce these proportions, the original datasets were trimmed from the end, shuffled based on a pre-selected window size (i.e., the number of sequential sample readings), concatenated based on type (e.g., slipping or non-slipping), and shuffled again based on a pre-selected window size. This process yielded approximately eight million data points (single sensor samples) for every individual sensor – half for training and half for testing the slip detectors.

E. Data Preprocessing

A unified approach was taken to produce a single, high-frequency data stream from a sensor's raw tactile data. Since a general force-sensitive tactile sensor may have multiple data outputs, a formula for obtaining a single data stream from all relevant sensor outputs involves calculating the lag-one finite difference of the l_2-norm of the sensor vector data,

$$s_t = ||\mathbf{a}_t||_2 - ||\mathbf{a}_{t-1}||_2 \tag{1}$$

where $s_t \in \mathbb{R}$ is the preprocessed sensor signal at timestep t; $\mathbf{a}_t, \mathbf{a}_{t-1} \in \mathbb{R}^m$ are the m-dimensional sensor vector data at timestep t and $t-1$, respectively; and $||\cdot||_2$ denotes the l_2-norm. This formulation effectively pools the full magnitude of the sensor's response to *changes* in force during tactile experiences, regardless of contact location. Furthermore, the lag-one difference produces a data stream centered around zero that signals changes in force loading while removing the absolute readings. The inspiration for this formulation comes from the fast-acting tactile afferents in the human hand that are only receptive to mechanical *transients*.

III. Spectral Analysis

While LSTM networks are used to perform slip detection, an in-depth spectral analysis of the gathered sensor data was conducted in order to characterize sensory response and provide insight into the feasibility of calibrating a sensor for slip detection. Effectively quantifying any frequency bands that yield statistically significant differences in sensory response between non-slipping and slipping data is useful in determining the feasibility for slip calibration and guiding sensor sampling rates.

After data balancing and preprocessing, the collected data for each sensor was pooled into one of two categories: non-slipping and slipping. Next, both non-slipping and slipping datasets were bootstrap-sampled 100 times, wherein each sample consisted of a sequence of sensor data of equal length to the maximum sampling rate of the sensor (e.g., 1000 data points per sequence for Sensor 1 and Sensor 2). Each non-slipping and slipping sequence was analyzed by the Fast Fourier Transform, yielding a single-sided amplitude spectrum at 1 Hz resolution over a frequency range of zero to one-half the maximum sensor sampling rate. The returned amplitudes per frequency for each set of bootstrapped samples were analyzed by the two-sample Kolmogorov-Smirnov (KS) algorithm between non-slipping and slipping data. The KS algorithm is a statistical, non-parametric distribution test

for continuous data that determines whether or not two sample sets belong to the same population [22]. By comparing 100 samples of non-slipping and slipping amplitudes per frequency, the KS algorithm exposed those frequencies for which the sensor data was the most distinguished between non-slipping and slipping events. This bootstrap sampling process and statistical test was repeated 200 times, yielding 200 KS tests per frequency. A single KS test either accepted or rejected the null hypothesis by returning a value of 0 or 1, respectively. Averaging the 200 KS test verdicts per frequency yielded a stable *significance* signal as shown in Fig. 2. Also shown are the mean and 95 % confidence bands of the amplitudes, indicating the variability in sensory response during non-slipping and slipping events.

There are a number of important results evident in Fig. 2. In particular, every sensor type exhibited contrasting features for non-slipping and slipping data. Sensors 1 and 2 yielded significant differences in signal amplitudes across the entire range of frequencies (0 Hz - 500 Hz). In contrast, Sensor 3 expressed two significant bands of frequencies including 0 Hz through 75 Hz and 175 Hz through 225 Hz. These frequencies fall directly within the sensing realm of FA-I afferents (5 Hz - 50 Hz) and FA-II afferents (40 Hz - 400 Hz). At approximately 200 Hz, both Sensors 1 and 3 experienced spikes in amplitudes for non-slipping data, the known idling vibration of the robotic arm. Sensor 2 appeared completely insensitive to the arm's idling vibration (likely due to the mechanical isolation and compliance of the sensor's air-filled rubber dome with only perimeter connections to the sensor base). Additionally, the most significant frequency band ($\sim$1 significance) that generally contains the largest separation in amplitude means is 0 Hz - 100 Hz for Sensors 1 and 2, and 0 Hz - 50 Hz for Sensor 3. Generally, both Sensors 1 and 2 experienced peak separation in mean curves at approximately 60-70 Hz, the dominant frequency of the catch-and-snap effect [5]. Unlike Sensor 1, the slip state bands for both Sensors 2 and 3 experienced increasing overlap beyond approximately 250 Hz. This effect was likely due to their thicker outer contacting rubber surfaces acting as mechanical low-pass filters to these very high frequency signals. Sensor 1 did not exhibit this behavior likely due to its higher overall rigidity and a rubber coating less than 0.2 mm thick.

Overall, these plots provided strong evidence that simply setting amplitude thresholds from FFT analyses of sensor data could result in a brittle slip detection strategy since major overlaps in amplitudes existed across the full frequency range between non-slipping and slipping data for all sensor types. Moreover, large data collection is still necessary to create these high-fidelity spectral plots to guide a thresholding technique. Regardless, such a technique was briefly investigated in Section V-A which confirmed a major loss in slip classification accuracy when compared with that of LSTM networks.

IV. Calibration for Slip Detection

Many existing approaches seek to extract relevant features from spectral algorithms that are either used directly to predict slip or to train a separate classifier to predict slip. Although these approaches have produced worthwhile results in the past, the feature extraction process may have failed to produce all relevant features in the data. Instead, this new approach passes minimally processed sensor data streams directly to the LSTM networks such that they are less likely to exclude obscured, yet relevant information for training a robust slip detector.

A. LSTM Architecture

The governing equations for the binary classification LSTM model (refer to [21], [23] for more details regarding LSTM cells) used for slip detection are:

$$z^t = g(W_z x^t + R_z y^{t-1} + b_z) \tag{2}$$

$$i^t = \sigma(W_i x^t + R_i y^{t-1} + b_i) \tag{3}$$

$$f^t = \sigma(W_f x^t + R_f y^{t-1} + b_f) \tag{4}$$

$$c^t = z^t \odot i^t + c^{t-1} \odot f^t \tag{5}$$

$$o^t = \sigma(W_o x^t + R_o y^{t-1} + b_o) \tag{6}$$

$$y^t = h(c^t) \odot o^t \tag{7}$$

$$y_p^t = \sigma(W_y y^t + b_y), \tag{8}$$

where W_z, W_i, W_f, $W_o \in \mathbb{R}^{N \times M}$ are the inputs weights; R_z, R_i, R_f, $R_o \in \mathbb{R}^{N \times N}$ are the recurrent weights; b_z, b_i, b_f, $b_o \in \mathbb{R}^{N \times 1}$ are the hidden bias weights; $W_y \in \mathbb{R}^{K \times N}$ is the output weights; and $b_y \in \mathbb{R}^{K \times 1}$ is the output bias weights. Furthermore, $x^t \in \mathbb{R}^{M \times 1}$ is the input vector; $z^t \in \mathbb{R}^{N \times 1}$ is the block input; $i^t \in \mathbb{R}^{N \times 1}$ is the input gate; $f^t \in \mathbb{R}^{N \times 1}$ is the forget gate; $c^t \in \mathbb{R}^{N \times 1}$ is the cell state; $o^t \in \mathbb{R}^{N \times 1}$ is the output gate; $y^t \in \mathbb{R}^{N \times 1}$ is the block output; and $y_p^t \in \mathbb{R}^{K \times 1}$ is the network output. Finally, $\sigma(\cdot)$ is the logistic sigmoid function; both $g(\cdot)$ and $h(\cdot)$ are the hyperbolic tangent function; and $\odot$ is the element-wise vector product. For binary classification, $K = 2$, and each output is a number between 0 and 1, which is interpreted as the likelihood of belonging to that output's class (i.e., either non-slipping or slipping). For a single, preprocessed sensory signal, $M = 1$. The number of hidden LSTM units, N, is determined below.

B. LSTM Network Parameters and Optimization

Through trial and error, the network parameters were fixed at values that produced relatively well-functioning slip detectors across all investigated sensors on the training datasets. The hidden layer size was fixed at 20 LSTM neurons, which yielded useful classification rates while minimizing training time and without overfitting. Momentum was held constant at 0.125. Mimicking a cooling schedule, four learning rates of 0.01, 0.001, 0.0001, and 0.00001 were applied sequentially for up to 100 training epochs, or until no measurable progress was demonstrated for 10 consecutive epochs. Optimization was conducted via full back-propagation through time (BPTT) by calculating the network gradients of a fully unfolded network. The typical

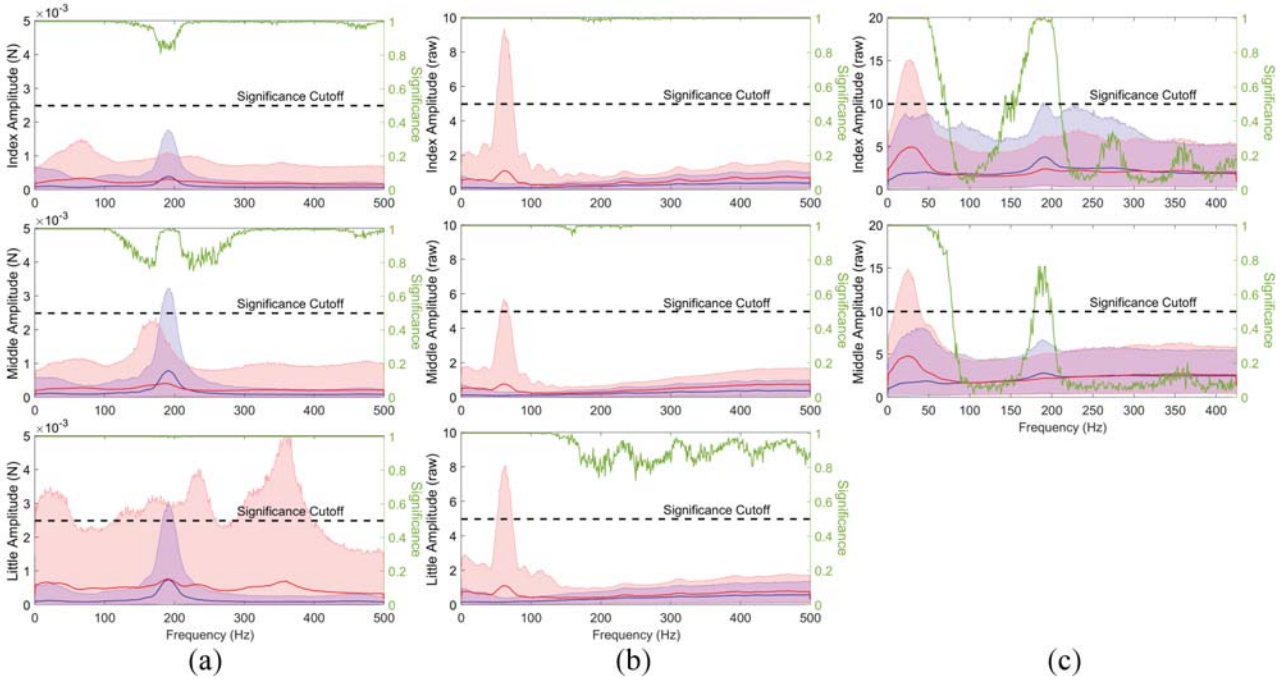

Fig. 2: Spectral plots of index and middle (a) Sensor 1, (b) Sensor 2, and (c) Sensor 3 indicating the mean amplitude (solid lines), 95 % confidence bands (faded areas), and significance (green line). Blue is non-slipping data, red is slipping data.

cross-entropy loss function was used for optimization. The input data stream to the networks were normalized by the 2.5^{th} and 97.5^{th} percentiles of the training datasets rather than the minimum and maximum values in order to reduce detrimental squashing of input data from the extreme outliers in the data. Training and testing models were conducted on a commodity laptop with a CPU core speed of 2.8 GHz. Training time took less than 23 μs per sensor sample, e.g., about 100 s per epoch for a dataset of 4.4 million sensor samples. Training time was invariant of model window size since scaling window size reciprocally scaled the number of unique sensor sequences from a fixed dataset (see Section V on window size). Inference took less than 5 μs per sensor sample, e.g., less than 0.3 ms for a window size of 50 samples.

Since this study investigated slip detection performance over a variety of factors and sensors, the network sizes were intentionally kept small to reduce computational cost. Although the subsequently trained networks were not fully optimized, they yielded reasonable performances and served as a point of comparison across the investigated sensors. For further network refinement, larger hidden layers and stacking LSTM layers could improve detection performance.

V. Sensor Parameters

For any tactile sensor, there are typically two controllable parameters of interest: window size and sampling rate. Window size dictates the number of sequential sensor readings that are passed to a classifier before a prediction is made. This parameter is particularly relevant for slip detection since stick-slip events across a tactile sensor are captured over a sequence of readings. Therefore, window sizes must be sufficiently large to capture the slip phenomenon. However, window size also directly inflates slip reaction time. As such, a minimized window size that affords high slip detection accuracy rates is crucial to improving slip reflexive behaviors during grasping or manipulation processes. Sensor sampling rates are often user-selectable and dictate how many sensor readings occur per second. High sampling rates are generally required for quality slip detection. However, minimizing sampling rates while preserving slip detection accuracy is pertinent to minimizing data acquisition resources and will assist tactile sensors with large numbers of sensory signals that cannot be sampled at excessively high sampling rates (e.g., pressure arrays) [6], [9], [13].

A. Window Size

LSTM networks were trained on the balanced datasets shuffled at six different window sizes – 5, 10, 25, 50, 100, and 200 – for data collected at the maximum sensor sampling rate. Overall, Sensors 1 and 2 exhibited the highest slip classification accuracies of over 90 % after 200 consecutive sensor samples were passed to the LSTM networks. In contrast, Sensor 3 yielded over 85 % slip classification rates[1]. The elevated classification performance for Sensors 1 and 2 (when compared to Sensor 3) correlated to both the larger disparities in slip state bands as discussed in Section III and their finer sensory resolution. Across all sensors, it is unclear whether the classification accuracies obtained from

[1]These classification rates are competitive with those of prior research. However, direct comparisons would be misleading since the datasets and manner of data collection are different.

TABLE I: Slip state classification accuracy at various window sizes across three types of sensors and fingers.

Sensor	Finger	Classification Accuracy (%) at Window Size					
		5	10	25	50	100	200
1	Index	74.7	80.9	85.5	88.0	89.6	90.4
	Middle	80.3	86.2	91.0	89.9	91.5	94.2
	Little	84.3	86.4	91.4	91.5	92.6	92.0
2	Index	80.3	86.0	90.3	91.3	91.0	91.7
	Middle	82.5	87.1	91.0	91.0	92.1	91.0
	Little	73.0	79.4	85.8	87.2	90.1	91.6
3	Index	66.4	74.8	82.5	86.5	86.1	86.2
	Middle	68.3	73.4	80.9	85.1	88.1	85.1

a window size above 50 samples yielded statistically significant improvements. However, accuracy generally decreased with window sizes below 50 samples. With as few as five samples, accuracy dropped by 15 % for Sensors 1 and 2 and 20 % for Sensor 3. Therefore, to decrease slip prediction time and maintain near-maximum accuracy, a window size of 50 samples yielding a slip detection time of 0.05 sec for Sensors 1 and 2 and 0.06 sec for Sensor 3 may prove adequate as these times are similar to human slip reflexes [3].

As a reference point, Holweg's thresholding technique [5], applied across Sensors 1-3 for a window size of 50 samples, yielded classification accuracies of 71.4 %, 80.9 %, and 76.4 %, respectively. Threshold values were determined in two steps: 1) the sum of signal amplitudes (from spectral analysis) over the most significant frequency band (identified in Section III) was calculated for every window of data in the non-slipping and slipping datasets; and 2) the location of largest separation in the empirical cumulative distribution functions of these sums from non-slipping and slipping data constituted the threshold value. Threshold values were 0.0015 N, 3.5 units, and 25 units for Sensors 1-3. Although simpler, this approach yielded classification accuracies that are 10 % to 20 % less than that of LSTMs.

B. Sampling Rate

The LSTM networks were trained with various data sampling rates to measure the impact on classification accuracy. Since the datasets can be sequentially downsampled by a factor of two by omitting every other sensor reading within a sample sequence, five additional sampling rates were investigated as shown in Table II. The datasets with a window size of 200 samples were chosen for downsampling such that the number of samples per sequence at the lowest sampling rate was greater than five (i.e., a sequence of appreciable length).

Classification accuracy is highly sensitive with respect to sampling rate. For Sensors 1 and 2, every increase in sampling rate by a factor of two yielded an approximate 2 % to 4 % gain in classification accuracy. This consistent and steady ascent in accuracy is explained by Fig. 2. Both Sensors 1 and 2 had significant disparities in spectral content between non-slipping and slipping data across the full range of available frequencies. Consequently, heightened sampling rates resulted in the continual acquisition of relevant features

TABLE II: Slip state classification accuracy at various sampling rates across three types of sensors and fingers. *Superscripts indicate associated sensor type for selected sampling rate.

Sensor	Finger	Classification Accuracy (%) at Sampling Rate				
		$31.25^{1,2}$ 26.56^{3}	$62.5^{1,2}$ 53.125^{3}	$125^{1,2}$ 106.25^{3}	$250^{1,2}$ 212.5^{3}	$500^{1,2}$ 425^{3}
1	Index	71.6	75.2	80.0	81.6	81.6
	Middle	75.8	79.7	82.2	86.6	88.6
	Little	80.0	85.1	89.1	87.6	89.6
2	Index	80.2	83.0	85.7	87.3	88.5
	Middle	83.1	85.8	87.8	89.0	89.4
	Little	71.4	74.7	78.2	81.0	83.8
3	Index	57.6	58.6	60.7	64.4	81.0
	Middle	58.1	59.9	63.1	66.7	78.4

at higher frequencies in the signal. Notably, the largest gains in accuracy were achieved within a sampling rate of 125 Hz. At this Nyquist frequency, signals with frequencies at approximately 62.5 Hz could be discerned without aliasing effects. Both Sensors 1 and 2 experienced the clearest disparity in signal amplitudes at precisely this frequency. Overall, Sensors 1 and 2 still yielded useful accuracies above 80 %, even with a factor of ten reduction in maximum sampling rate (i.e., at 125 Hz). This indicated that useful slip detection accuracies were still tenable at relatively low frequencies, an attractive trait for tactile sensors with a large number of sensory outputs.

For Sensor 3, significant increases in accuracy were achieved when moving from a sampling rate of 250 Hz to 500 Hz. Referring to Fig. 2, Sensor 3 yielded a significant difference in spectral content between 175 Hz and 225 Hz. With a Nyquist frequency of 500 Hz, signals with frequencies up to 250 Hz could be seen without aliasing effects. However, at a sampling rate of 250 Hz, the significant signal content within 175 Hz and 225 Hz led to the approximate 15 % loss in classification accuracy. Again, steady increases in classification accuracy of 2 % to 4 % were sustained for every increase in sampling rate by a factor of two. This trend is logical, considering the first significant band of frequencies spanned 0 Hz to 75 Hz. In contrast to Sensors 1 and 2, Sensor 3 required sampling at rates greater than 500 Hz in order to achieve reasonable slip detection accuracies.

VI. Exogenous Effects

A truly robust tactile sensor should yield minimally varying slip detection accuracies regardless of material and slip speed effects. However, different materials possess differing tribological properties and slip speed is known to positively correlate with increasing signal amplitude [2]. Therefore, the previously calibrated slip detectors trained with a window size of 50 samples and at the sensors' maximum sampling rate (settings determined to yield maximum classification accuracy with quickest detection time) were tested against these factors.

TABLE III: Slip state classification accuracy with five different materials across three types of sensors and fingers.

Sensor	Finger	Classification Accuracy (%) with Material				
		Aluminum	PVC	Neoprene	Cardboard	Plywood
1	Index	90.9	87.1	76.1	91.1	81.2
	Middle	92.8	88.6	79.6	92.2	90.9
	Little	90.7	89.3	86.0	90.4	91.4
2	Index	86.5	89.5	91.9	77.0	86.9
	Middle	93.7	92.7	92.1	67.9	90.0
	Little	90.5	86.0	88.9	86.6	80.6
3	Index	81.1	83.2	84.6	79.3	79.6
	Middle	81.1	80.8	80.9	79.4	81.1

A. Material

Both slipping data collected at the four speed settings and non-slipping data for each of the five material types were pooled per material. The trained slip detectors were tested on these material-centric datasets to measure material sensitivity (see Table III). Sensor 1 experienced relatively minor fluctuations in slip classification accuracy among all material types except for neoprene, which had multiple accuracy rates below 80 %. Thus, Sensor 1 less accurately predicted slip on a softer material. Similarly, Sensor 2 was mostly invariant to material type except for cardboard, which had multiple accuracy rates below 80 %. During data collection, Sensor 2 noticeably gathered cardboard fibers on its rubberized contact surfaces, which likely negatively impacted its tactile sensations for predicting slip. Sensor 3 appeared to be the most invariant to material type, with small fluctuations in accuracy around a nominal of 80 %. Overall, slip detection accuracy was fairly consistent across all three sensor types. Collecting data from a greater variety of softer materials may improve the classification performance of Sensor 1. Periodic cleaning of outer surfaces of tactile sensors may also improve the classification performance of Sensor 2 with shedding materials.

B. Slip Speed

Both non-slipping and slipping data collected across the five materials were pooled for each of the four arm speed settings. The previously trained slip detectors were tested on these speed-centric datasets to measure their performance sensitivity to slip speed effects. Table IV illustrates that the greatest loss in slip classification accuracy across all sensor types was experienced at the smallest slip speeds (5 mm/s). Classification accuracy remained relatively consistent at the remaining three arm speeds. This result confirms that classification accuracy is indeed a function of slip speed. Specifically, accuracy increases with increasing slip speed. Overall, Sensor 1 was the most robust to slip speed effects, with slip classification accuracies above 80 %. Sensor 2 behaved similarly with the exception of one sensor, which achieved a classification accuracy below 80 %. Sensor 3 was the most sensitive to slip speed with a classification accuracy as low as 67 %.

TABLE IV: Slip state classification accuracy at four arm speed settings across three types of sensors and fingers.

Sensor	Finger	Classification Accuracy (%) at Speed			
		5 mm/s	25 mm/s	50 mm/s	75 mm/s
1	Index	80.8	86.7	86.6	87.1
	Middle	82.6	89.8	91.0	91.6
	Little	82.5	90.6	91.7	92.5
2	Index	80.1	87.6	88.8	88.9
	Middle	82.6	87.8	88.8	89.2
	Little	75.3	88.6	91.1	91.5
3	Index	71.4	84.1	85.5	85.2
	Middle	67.6	83.8	86.1	86.1

VII. Manufacturing Variability

The effect of manufacturing variability directly impacts the cost of calibrating a sensor for robust slip detection. Calibration costs are drastically reduced if the calibration models of one sensor can be directly re-applied to another sensor of the same type without any additional data collection. To quantify the impact of manufacturing variability, a slip detection model trained using data gathered from a sensor on one finger was applied to data of the same sensor type on another finger. The models used were again those previously trained with a window size of 50 samples and at the sensors' maximum sampling rate. Additionally, a new model was trained per sensor type on data acquired from all instances of that sensor type (a "combined" dataset). Evaluating this model on both the individual and combined datasets provided insight on performance gains from batching sensory data acquired across multiple instances of the same sensor type. Tables V-VII show the results for all sensor types.

Encouragingly, all sensor types exhibited a relatively high level of model transference. Loss in classification accuracy remained below 11 % for Sensor 1, 21 % for Sensor 2, and 3 % for Sensor 3. This result appears to indicate that Sensor 3 had the highest level of model transference, followed by Sensor 1 and Sensor 2. However, confidence in this assessment concerning Sensor 3 is diminished since only two sensors of this type were available. Regardless, training on data collected by all instances of a particular sensor type facilitated the equalization of the classification accuracy. Sensors 1, 2, and 3 exhibited a classification difference within 5 %, 3 %, and 2 %, respectively, among all sensor instances. Conclusively, immediate model transfer among sensors within a sensor type is possible. However, classification accuracy can be further equalized by training with data collected by a batch of sensors. This action will help reduce performance bias towards a particular sensor, and will likely generalize the model more accurately to all future instances of sensors for that particular type.

VIII. Conclusion

Slip detection as a sensing modality for tactile sensors was motivated by existing neurophysiological research of human hands. Extensive data collection captured the response signals of three completely different, commercial tactile sensors across various factors including slip speed, material type,

TABLE V: Slip classification accuracy for Sensor 1 with LSTMs trained and tested on data from one or all fingers.

		Testing Data			
		Index	Middle	Little	Combined
Training Data	Index	88.0	89.5	84.7	87.5
	Middle	82.5	89.9	86.1	86.2
	Little	80.5	83.7	91.5	85.2
	Combined	84.8	89.1	89.1	87.8

TABLE VI: Slip classification accuracy for Sensor 2 with LSTMs trained and tested on data from one or all fingers.

		Testing Data			
		Index	Middle	Little	Combined
Training Data	Index	91.3	89.9	73.0	84.8
	Middle	87.7	91.0	70.4	83.2
	Little	85.3	87.6	87.2	86.7
	Combined	91.5	92.2	89.4	91.0

TABLE VII: Slip classification accuracy for Sensor 3 with LSTMs trained and tested on data from one or all fingers.

		Testing Data		
		Index	Middle	Combined
Training Data	Index	86.5	83.5	85.1
	Middle	86.0	85.1	85.7
	Combined	86.9	85.4	86.3

contact force, and free-space motion. Spectral analysis and statistical testing of captured data yielded valuable insight into sensor behavior such as highlighting the most significant frequency bands for discriminating non-slipping and slipping events. LSTM networks were trained to analyze sequences of sensor data to predict the state of slip. The performance measurement methodology uncovered that larger window sizes and higher sampling frequencies improved slip detection accuracies. Moreover, sensors exhibited robustness to slip speed, material types, and manufacturing variability with the exception of a few edge cases. Future work includes injecting the trained slip detectors in a grasp reflex and quantifying the performance thereof.

REFERENCES

[1] 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, p. 345, 2009.

[2] R. D. Howe and M. R. Cutkosky, "Sensing skin acceleration for slip and texture perception," in *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on*. IEEE, 1989, pp. 145–150.

[3] R. S. Johansson and G. Westling, "Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects," in *Experimental Brain Research*, vol. 56, no. 3, 1984, pp. 550–564.

[4] Y. Yamada and M. R. Cutkosky, "Tactile sensor with 3-axis force and vibration sensing functions and its application to detect rotational slip," in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*. IEEE, 1994, pp. 3550–3557.

[5] E. Holweg, H. Hoeve, W. Jongkind, L. Marconi, C. Melchiorri, and C. Bonivento, "Slip detection by tactile sensors: Algorithms and experimental results," in *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, vol. 4. IEEE, 1996, pp. 3234–3239.

[6] J. A. Alcazar and L. G. Barajas, "Estimating object grasp sliding via pressure array sensing," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1740–1746.

[7] J. S. Son, E. A. Monteverde, and R. D. Howe, "A tactile sensor for localizing transient events in manipulation," in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*. IEEE, 1994, pp. 471–476.

[8] D. Goger, N. Gorges, and H. Worn, "Tactile sensing for an anthropomorphic robotic hand: Hardware and signal processing," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 895–901.

[9] B. Heyneman and M. R. Cutkosky, "Slip classification for dynamic tactile array sensors," *The International Journal of Robotics Research*, vol. 35, no. 4, pp. 404–421, 2016.

[10] W. Yuan, R. Li, M. A. Srinivasan, and E. H. Adelson, "Measurement of shear and slip with a gelsight tactile sensor," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 304–311.

[11] M. Vatani, E. D. Engeberg, and J.-W. Choi, "Force and slip detection with direct-write compliant tactile sensors using multi-walled carbon nanotube/polymer composites," *Sensors and Actuators A: physical*, vol. 195, pp. 90–97, 2013.

[12] Z. Su, K. Hausman, Y. Chebotar, A. Molchanov, G. E. Loeb, G. S. Sukhatme, and S. Schaal, "Force estimation and slip detection/classification for grip control using a biomimetic tactile sensor," in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*. IEEE, 2015, pp. 297–303.

[13] R. Fernandez, I. Payo, A. S. Vazquez, and J. Becedas, "Micro-vibration-based slip detection in tactile force sensors," *Sensors*, vol. 14, no. 1, pp. 709–730, 2014.

[14] J. Reinecke, A. Dietrich, F. Schmidt, and M. Chalon, "Experimental comparison of slip detection strategies by tactile sensing with the biotac® on the dlr hand arm system," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 2742–2748.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[16] G. Melis, C. Dyer, and P. Blunsom, "On the state of the art of evaluation in neural language models," *arXiv preprint arXiv:1707.05589*, 2017.

[17] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey *et al.*, "Google's neural machine translation system: Bridging the gap between human and machine translation," *arXiv preprint arXiv:1609.08144*, 2016.

[18] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[19] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, and D. Quillen, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *The International Journal of Robotics Research*, p. 0278364917710318, 2016.

[20] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio," *arXiv preprint arXiv:1609.03499*, 2016.

[21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[22] M. A. Stephens, "Edf statistics for goodness of fit and some comparisons," *Journal of the American statistical Association*, vol. 69, no. 347, pp. 730–737, 1974.

[23] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "Lstm: A search space odyssey," *IEEE transactions on neural networks and learning systems*, 2017.