2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)
June 28-30, 2023. Seattle, Washington, USA

Spectro-Temporal Recurrent Neural Network for Robotic Slip Detection with Piezoelectric Tactile Sensor

Théo Ayral[1,2], Saifeddine Aloui[1], and Mathieu Grossard[2]

***Abstract*—In this paper, we present a novel tactile-based method for detecting slippage in robotic manipulation, using a single piezoelectric sensor. The method combines spectral analysis (FFT) and deep learning (GRU) for improved efficiency and adaptability. We implement an automated data-collection process with accurate and unbiased labels of slip events. The proposed method is evaluated through an ablation study characterizing the influence of model hyperparameters and interaction settings. The results show a high classification accuracy of 98.70% at 100Hz and detection delays of 8.5 ± 23.7ms, demonstrating the relevance of our spectro-temporal pipeline. The proposed method has the potential to enhance the performance of robotic systems and increase their reliability in robotic grasping applications.**

I. INTRODUCTION

Context

Robotic manipulation tasks, in which objects need to be physically interacted with, involve contact interactions that are best monitored through tactile sensing. Slip detection is a critical aspect in robotic manipulation, as it can prevent costly mistakes and ensure the successful completion of tasks. The term "slip" refers to relative motion between two surfaces in contact, notably in robotic manipulation, between the robotic fingers and the grasped object. Even small slippage induces loss of information about object pose and location, which can lead to task failure [1]. The naive approach to prevent slip is applying maximal tightening force, thus increasing friction to ensure grasp stability. By detecting slip at an early stage, corrective actions can be taken to prevent substantial slippage, allowing the robotic finger to minimize the force and energy required to maintain stable grasp control, and to prevent the fingers from damaging the object [2]. Furthermore, slippage could be understood more generally as a specific type of tactile event, involved in many dexterous tasks (such as tactile exploration or human-robot interactions).

Related Work

Friction plays a major role in grasp stability, and is the foundation of force-closure [3]. Thus, most slip detection methods are based on the analysis of the friction phenomenon, either through analytical models [4] or with data-driven approaches [5]. Analytical models often derive from the Coulomb friction law and are interested in friction coefficients and tangential-to-normal force ratios. Alternatively, dynamic mechanical signals at the contact interface can be detected and processed to identify friction vibrations [6], for gross-slip or incipient-slip detection. Data-driven approaches can use a broader range of sensory input from which a learning model can extract characteristic features of slippage [7].

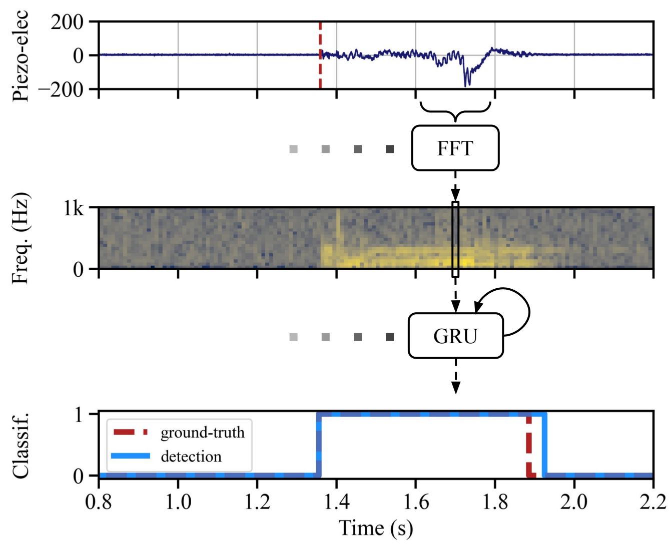

Fig. 1. FFT-GRU pipeline for slip detection. The piezoelectric sensor captures mechanical stress due to friction vibration. FFT extracts spectral features that are processed by the recurrent network. Contact mode is classified as *hold / slip* at 100Hz.

1) Friction-vibration detection: Our approach for slip detection is based on the force dynamics at contact interface, captured by a piezoelectric sensor [8]. Comparative studies of sensor technologies can be found in relevant reviews [1]. Inspiration comes from the human hand's Fast Adaptive mechanoreceptors which detect vibrations during slippage [9] [10]. The piezoelectric sensor reacts to dynamic events, generating electric charges in response to changes in the stress field within the material. Performing a frequency-domain analysis of the dynamic signal enables identification of a spectral profile that is characteristic of friction vibrations. Relevant frequencies can be selectively monitored to inhibit signals coming from motor vibrations and other sources [7]. While this line of work relies on the assumption of a characteristic frequency pattern of slippage dynamics, research does not seem to converge to a simple solution and the vibration profile is dependent on a number of parameters of the environment, the robot sensor, and the interaction. Reported relevant frequencies for slip detection can range from 40Hz to 1kHz [11] [7] [2]. This again

*This research was supported by TraceBot project. TraceBot has received funding from the European Union's H2020-EU.2.1.1. INDUSTRIAL LEADERSHIP programme (grant agreement No 101017089).

[1] CEA, Leti, 38000, Grenoble, France.
{theo.ayral, saifeddine.aloui}@cea.fr
[2] Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France.
{theo.ayral, mathieu.grossard}@cea.fr

makes learning methods particularly relevant as they can extract features and find patterns for a wide range of setups, provided that training data has sufficient variability. In the human hand, distribution of fast-adaptive mechano-receptors FA-I and FA-II suggests the 5-50Hz band is relevant for tactile exploration with the fingertips, while 40-400Hz provides information on grasp, through the palm, for object-holding tasks [9].

2) Learning models: Typical machine-learning classifiers as Support Vector Machine and Random Forest have already shown good results for slip detection, with generalization to novel objects [12] [13]. Deep learning models are also very common. Temporal Convolutional neural Networks (TCN) and Recurent Neural Networks (RNN), notably Long Short-Term Memory (LSTM), process data in a structured way that is well-adapted to time series such as tactile-sensor signals [5] [14] [15] [16]. Despite these advances, collecting large training data remains a major challenge for generalizing beyond constrained environments and capturing the diversity of possible manipulation scenarios in real-world robotics.

3) Incipient slip and slip prediction: Mandil *et al.* proposed a slip-prediction system by classifying predicted future sensor signals [13]. This is a very complex problem and it requires taking future robot actions into account, as they influence tactile states. Incipient-slip detection, on the other hand, focuses the phenomenon of local slippage, with micro-displacements already happening. This behaviour can be favoured directly by the hardware design [17] [10]. In this work, the objective is to perceive the actual contact state, considering that slippage is not inherently something to avoid, but an integral part of dexterous manipulation [4] [16].

Contributions

We study the capabilities of a learning model to detect slip with a single piezoelectric sensor, through a combination of spectral and temporal analysis. Our proposed detection pipeline leverages the efficiency of the Fast Fourier Transform (FFT) to extract relevant representations of the friction vibration, and the flexibility of deep learning to generalize across a wide range of slip-event parameters, such as speed, texture and force. Gated Recurrent Units (GRU) offer a mechanism to extract data-driven temporal patterns that are shown to be relevant to the task of slip detection. An overview of the pipeline is presented in Fig. 1. Instead of collecting data during an actual manipulation task, which can be laborious and inaccurate, we setup an automatic data-collection bench with parameterized slip-trajectory generation and unbiased labels.

Section II briefly describes the design of our piezoelectric tactile sensor. Section III discusses the formulation of the detection task and delves into the data-collection process. Section IV details our spectro-temporal model, its motivation and implementation. In Section V, various ablations of the model are conducted, and its performance is rigorously evaluated. Finally, Section VI summarizes our findings and concludes the paper.

II. Piezoelectric tactile sensor

The piezoelectric sensor is embedded underneath a flexible mechanical support shown in Fig. 2. Ten piezoelectric generator layers are stacked for better energy harvesting and performance. They constitute alternating layers of electrodes (800nm thick high-conductivity polymer, PEDOT-PSS) and electroactive polymer (2.5µm thick ferroelectric ink, copolymers PVDF-TrFE). The sensor spans 20×10 mm^2 at the center of the tactile area. The electronic board samples the piezoelectric signal with a 10kHz frequency.

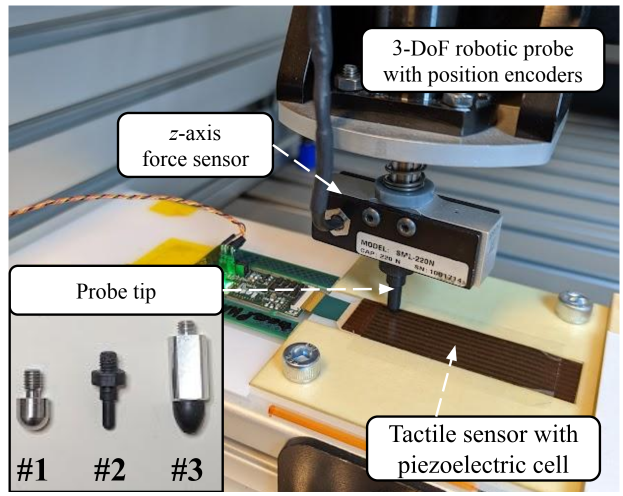

Fig. 2. Robotic data-collection setup with 3D-position and 1D-force ground-truth. The tactile sensor embedding a piezoelectric cell is fixed on the bench and touched by a mechanical probe. Tips with different shape, texture and softness provide variability in the generated slips.

III. Data collection

Obtaining accurate ground-truth information on the slippage of un-instrumented objects presents a significant challenge.In order to not only evaluate but also train the slip detector with labeled data, we implement a fully-automated data collection and labeling system.

A. Methodology

Slip detection systems in robotics research are often evaluated on simple manipulation tasks, using grasp stability metrics as a high-level proxy for slip-detection performance. Actual validation with accuracy and delay of slip detection is hard to perform but is still necessary [18]. As underlined by Waters *et al.*, grasp success as a metric can be relevant for slip detection only if the mechanism of grasp force minimization is also characterized [17]. Using conservative slip detection ensures grasp stability easily but can miss the point of energy minimization [19] [12]. Proper validation should comprise variations of environment and task parameters, such as object material and geometry, slip speed, grasp force [10].

In addition, data-collection requires labels. In the work of Yan *et al.*, a human annotator labels slippage manually by observing the grasp [20]. Zhang *et al.* invoke “the intuitive human sense of slip” as the best alternative to an objective labelling criterion, which is difficult to implement [14]. Annotation can also be performed in video recordings [12] and the process can be automated with visual detection of markers [21]. Frame-rate and image resolution are limiting factors for the application of vision-based method for accurate labeling of slip event, in space and time dimensions.

To overcome this challenge, we collect sensor data jointly with ground-truth signals by using an automated characterization test bench instead of relying on the target manipulation platform. The robotic bench (Fig. 2) can move an end-effector in space, following straight trajectories. A 1D reference force sensor (PMI SML-220N) monitors force in the normal axis z. We generate trajectories with random positions, speed and force in contact with the sensor to collect slippage data. Recorded position measurements sampled at 100Hz are used to obtain ground-truth slippage timings, as illustrated in Fig. 3. This approach involves stepping out of the traditional manipulation framework as we measure slippage between the tactile sensor and an instrumented robotic probe. This setup combines the benefits of instrumentation and automation, allowing for fully-automated collection of rich slippage data.

B. Dataset

In addition to slip segments, sensor data is recorded when the probe moves in the air, without contact. The tactile sensor captures vibration noise caused by the robot motors (see Fig. 5). We also collect slip segments with three end-effector tips (Fig. 2), of different texture shape and compliance, to provide variability in friction coefficient and force distribution in the recorded slips. The dataset contains 3,200 recordings with 1,600 containing a slip segment, and 1,600 with motor noise only. The average duration of recordings is 3.33s, and 1.5s for slip segments. Slip trajectories are generated with random parameters in the range 2-10N for contact force, 10-32mm for travel distance, 200-2000 mm/min for speed.

IV. Hybrid Spectro-Temporal Classification Algorithm Rationale

The pipeline is composed of spectral analysis of the piezoelectric signal, followed by learned temporal-feature extraction and binary classification. An overview is proposed in Fig. 1.

A. Spectrogram

We compute Power Spectral Density (PSD) with Fast Fourier transforms (FFT) on the mean-subtracted piezoelectric signal, with temporal windows of 200 values (20ms at 10kHz). The spectrogram of the entire signal is built by concatenating the sequence of FFTs, forming a 2D array with frequency and time axes. It contains information on the frequency content of the signal, and its evolution over time (see Fig. 3). We compute FFTs at intervals of 10ms, with an input window size of 20ms. This results in an overlap between adjacent FFTs in the spectrogram. Providing longer segments as FFT input improves frequency resolution. The frequency bins are in range 0-5kHz with a resolution of 50Hz per bin.

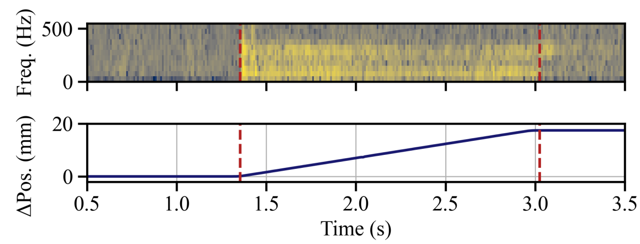

Fig. 3. Slip segment example, with piezoelectric signal spectrogram and ground-truth position measurements providing slip timing labels.

B. Stacked Gated Recurrent Units

The resulting spectrogram is then fed to a Recurrent Neural Network (RNN), processing one FFT at a time, to detect temporal patterns. Classification is performed on each time-step, every 10ms. Our limitation is the label frequency, due to 100Hz sampling of the ground-truth position sensor.

1) Architecture: We train a recurrent neural network of 4 stacked Gated Recurrent Units (GRU) [22]. It produces a 32-channel representation for each time-step ("sequence-to-sequence"). These outputs are then classified independently by a fully-connected linear layer. The network produces a temporally dense classification output. The temporal piezoelectric signal is transformed into a sequence of FFTs (spectrogram), which is then classified into a sequence of binary class predictions (*hold* or *slip*). The model's architecture is illustrated in Fig. 4, highlighting the sequential processing.

2) Training: RNN and classifier are trained jointly end-to-end. We use RMSProp optimizer with learning rate $8e^{-4}$ and Cross-Entropy loss. The network is trained for 200 epochs, keeping the model-state which obtained the best validation performance (balanced accuracy). In order to build training mini-batches (size 32), we extract short random clips of 128

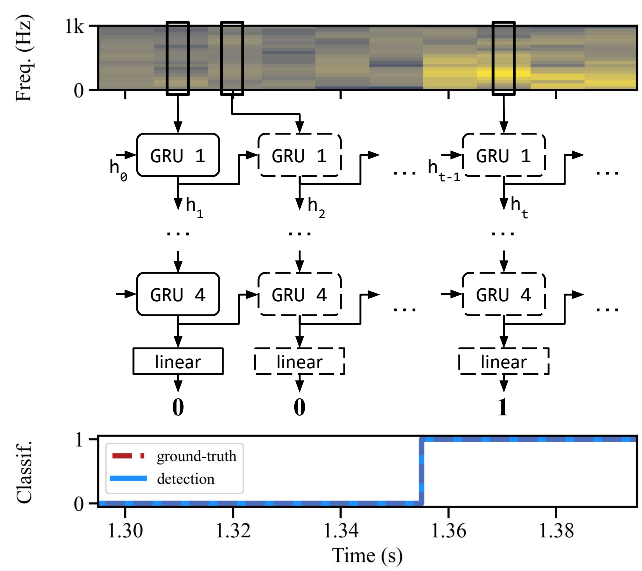

Fig. 4. Model architecture based on four stacked Gated Recurrent Units (GRU) classifying frequency-domain features to detect slip. The network processes the signal sequentially for a live scenario application, with 100Hz classification.

frames from the spectrograms. This is also a form of data-augmentation as the RNN processes the signal in the temporal direction, and the memory is initialized before processing a clip. By cropping the signal randomly, we vary the quantity of past information provided for the RNN during training.

We use a smooth-change penalty (L_{scp}) to favour temporally smooth predictions in parts where the label does not change. This loss is only activated after the 20th epoch to allow more stable training in early epochs. It is computed as the sum of the absolute first-order temporal difference of y_t, the 2-channel output logits :

$$L_{scp} = \sum_{t=1}^{T} |y_t - y_{t-1}|. \tag{1}$$

For additional regularization, we use label smoothing on clips containing slip events, because noise in the labels could impede model training.

The dataset has unequal numbers of frame-level samples per class. A working solution is using class weights in the loss function. However, we found that a balance also has to be ensured within the non-slip frames, as both pre-slip and post-slip frames possibly have different dynamics. This also means the model needs to be trained to detect starting of slippage but also end of slippage. The RNN has to consider end of slip as a reset of its hidden state, so as to be ready to detect slippage again. Our solution is to enforce balance in the training batches directly with probability distributions when sampling short clips from the recordings. A common concern of slip detection in robotics is the confusion between motor noise and slippage vibrations [10]. Training and testing with noise data, as described in Section III-B, ensures the model does not leverage motor vibration as an indication of possible slippage, but also successfully extracts slippage features from noisy signals.

C. Discussion

This pipeline involves a hierarchy of dynamic feature extraction. The piezoelectric sensor inherently provides a measurement of dynamic changes in mechanical deformation. Then, FFT computes the frequency components in the signal. Finally, the GRU identifies temporal patterns in the evolution of this spectral profile. This results in a high-level and robust analysis of the contact dynamics.

A typical alternative to RNNs in such tasks is Temporal Convolutional Networks (TCN). They can be preferred for their better capacity to retain past information and detect long-term dependencies [15]. TCNs have direct access to past information while the RNN only stores a hidden state that is supposed to summarize information and extract relevant context. In fact, this mechanism seems well suited to the slip detection process, in which precise long-range dependencies are not relevant, but we are more interested in short-term patterns and past context. Also, the sequential processing of RNN is a constraint when training, as it limits parallel computation. But in live scenarios, it is very practical as data are received and processed in a sequential way, and robot controllers do not usually run on GPUs.

V. Experimental characterization and results

We perform various ablations on the model to assess the importance and impact of each component in the proposed slip-detection pipeline. We provide the balanced binary accuracy of *hold / slip* contact classification at frame-level. To have more in-depth understanding of the system's behavior, we also report F_1-score on binary classification for slip detection [12], combining both precision and recall metrics: $F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$.

We perform k-fold cross-validation with $k = 5$. The dataset is composed of 3,200 recordings, half of which contain slip while others only record motor noise. Unless specified otherwise, the following experiments are performed with the FFT-GRU pipeline, with feature dimension 32, using training clips of 1.28s and frequency range 50-600Hz.

A. Architecture ablation

Table I reports model performance when evaluating different architectures, with ablation to the proposed spectro-temporal pipeline. Hyperparameter *dim* specifies the feature dimension of the model while *ker* is the kernel size for the convolutional models. We report the total number of parameters for each model. First, we verify that spectral features provide good class separability for our task. Van Wyk and Falco showed that RNN can work directly on a combination of tactile-sensor signals [5], but this approach does not fit well with our high-frequency piezoelectric sensor. Computing the FFT explicitly provides an efficient way to extract discriminative features, reducing the complexity of the model. Our FFT-GRU model obtains 98.70% balanced accuracy. Removing the FFT step and feeding the raw piezoelectric sensor signal requires down-sampling to 100Hz, as we are limited by ground-truth sensor frequency. This results in a drop to 94.71% when the network works directly on the raw piezoelectric signal (*raw pE*).

TABLE I
ABLATION ON SPECTRAL AND TEMPORAL FEATURE-EXTRACTION

Input	Model	*dim*	*ker*	Params	F_1-score	Acc. (%)
raw pE	GRU	32	–	22k	0.9157	94.71
FFT	GRU	32	–	23k	0.9781	98.70
FFT	GRU	128	–	352k	**0.9787**	**98.73**
FFT	MLP	32	–	3.6k	0.7913	88.76
FFT	MLP	128	–	51k	0.7941	89.39
FFT	TCN	128	1	51k	0.8051	88.97
FFT	TCN	128	8	405k	0.9579	97.64
FFT	TCN	128	32	1.6M	0.9694	98.26
"	"	"	"	"	0.9836*	99.21*

*Non-causal model.

We also validate the relevance of temporal patterns in the spectrogram, by experimenting with a single-frame Multi-Layer Perceptron (MLP), classifying each FFT independently. The accuracy does not exceed 90%, even when augmenting the network's dimension. To better understand the effect of temporal features, we experiment with Temporal Convolutional neural

Networks (TCN). We can change the temporal receptive field of the network through the convolution-kernel size. To simplify the analysis we use a very naive implementation by stacking 1D-convolutional layers, making model parameters linearly increase with receptive field. When $ker = 1$, it is equivalent to the single-frame MLP. A more efficient approach to increase the temporal receptive field would use temporal pooling and kernel dilation [15]. The reported TCN performances are not aimed for a comparison with the RNN approach, as number of parameters and non-linearities do not scale identically for both network architectures. The results clearly show performance increasing with the temporal receptive field, from 88.97% when working on a single 20ms FFT, to 98.26% when working on 320ms spectrogram windows.

As the objective is real-time slip-detection, causality has to be enforced so the model does not have access to future data. This is straightforward with the RNN but requires padding the input of TCN.

B. *Informative frequency range*

The spectrogram has a frequency range of 5kHz due to computing FFT on a 10kHz piezoelectric signal. The 50Hz bin resolution is determined by the duration of the provided segments (20ms in our setup). Literature and data visualization suggest friction vibrations might be limited to certain frequencies (see Section I-1). Feeding the full spectrogram's frequency features to the model could result in waste of computation. Thus, we experiment with different frequency ranges as input to the GRU stack (Table II).

TABLE II
ABLATION ON SPECTROGRAM FREQUENCY RANGE

Freq. range (Hz)	F_1-score	Acc. (%)
0 – 5000	0.9589	97.40
0 – 600	0.9777	98.67
50 – 400	0.9769	98.40
50 – 450	**0.9784**	**98.80**
50 – 500	0.9783	98.79
50 – 600	0.9782	98.74*
50 – 1000	0.9774	98.63

*This frequency range is used in other experiments.

We find that slip-related vibrations can start as low as 50Hz and range up to about 450Hz. While higher frequency bands could be affected by slip vibrations, this information might be redundant. These results are coherent with previous literature on robotic manipulation and human touch, as discussed in Section I-1. In addition, reducing the spectrogram to relevant frequencies is beneficial to network performance, as it can facilitate training, because of reduced computations and better signal-to-noise ratio.

C. *Generalization to Unseen Shape and Material*

Data was collected using three interchangeable probe tips with different texture and shape (as displayed in Fig. 2). To investigate the generalization capabilities of the model, we perform a Leave-One-Subject-Out (LOSO) evaluation.

TABLE III
GENERALIZATION TO UNSEEN SHAPE AND TEXTURE

Left-out end effector	F_1-score	Acc. (%)
None	0.9762	98.60
#1	0.9502	96.45
#2	0.9370	95.32
#3	0.9337	94.41

The variation in shape is actually very limited and interactions are notably constrained to soft-finger contacts, with small area of contact. Yet, even in this framework, the model does not display good generalization capabilities (see Table III). We consider this is a limitation of the data. When performing the LOSO procedure, there are only two different tips in the training data, providing insufficient variability.

TABLE IV
PERFORMANCE WITH AND WITHOUT NO-SLIP MOTOR-NOISE SAMPLES

		Training dataset			
		Slip		Slip + Noise	
Test	Slip	**F_1-score** 0.9831	**Acc.** 98.59%	**F_1-score** 0.9766	**Acc.** 98.16%
	Slip + Noise	**F_1-score** 0.6779	**Acc.** 86.50%	**F_1-score** 0.9735	**Acc.** 98.53%

D. *Immunity to Motor Noise*

Tactile signals recorded during free motion without contact, labelled as non-slipping, are very effective for training the model to dissociate friction vibration from motor noise. An example is provided in Fig. 5. However, the generalization to different robots and motor-vibration profiles requires further investigation we leave for future work. Simulated noise is another approach we expect to be relevant for this issue, but our setup does not allow to validate it for now.

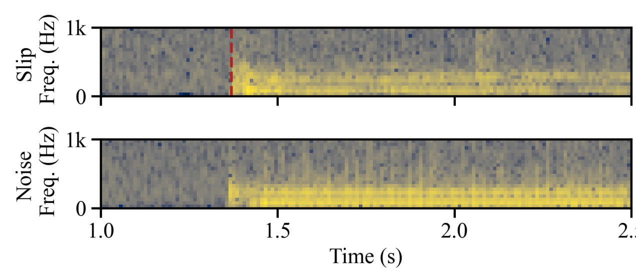

Fig. 5. Slip signal (Top) and motor noise sample without slipping contact at the sensor interface (Bottom). Both spectrograms have the same normalization for comparison.

As shown in Table IV, the motor vibrations are misclassified as slip signal if such data is not provided for training explicitly. However, when considering motor noise, the classification task becomes more difficult and the model suffers a small performance loss on classifying slip segments, with -0.43%.

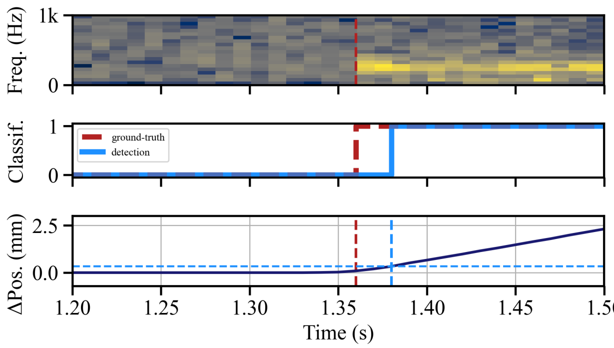

Fig. 6. Slip detection delay. The model detects slippage 20ms after the ground-truth. The total traveled distance during this time is 0.8mm.

E. Results : Detection delay

Beyond frame-level classification accuracy, slip detection is interested in the detection delays, computed as the time interval between ground-truth slip start and slip detection by the FFT-GRU model. On 3,200 samples, 7 were incorrectly recognized in a way that prevents computing a detection delay. For the others, the average detection delay was 8.51ms, with 23.7ms standard deviation. This result has to be considered with an error margin of 10ms due to the ground-truth signal sampling-rate. A visualization is provided in Fig. 6. Short detection delays make the method well-suited for grasp stabilization in manipulation tasks [23]. The maximum delay was 58ms, which is very similar to the result reported in the work of Van Wyk and Falco [5].

Motor noise frames are correctly classified at 99.85%. On the slip recordings, balanced frame-level accuracy is 98.24%, with an F_1-score of 0.978 (precision : 0.973, recall : 0.983). When considering both slip and noise segments, the balanced accuracy is 98.70%.

VI. CONCLUSIONS

We have demonstrated the possibility of detecting slippage with a high-frequency piezoelectric sensor embedded beneath a robotic tactile skin. Taking advantage of the efficiency of spectral representations for discriminative features, and leveraging data-driven models for capturing temporal patterns, we designed a classification pipeline combining Fast Fourier Transforms and Gated Recurrent Units. The proposed detection pipeline was kept light-weight to operate at 100Hz and showed quick detection for real-time applicability to manipulation tasks.

REFERENCES

[1] M. Park, B.-G. Bok, J.-H. Ahn, and M.-S. Kim, "Recent Advances in Tactile Sensing Technology," *Micromachines*, vol. 9, no. 7, p. E321, Jun. 2018.

[2] R. A. Romeo and L. Zollo, "Methods and Sensors for Slip Detection in Robotics: A Survey," *IEEE Access*, vol. 8, pp. 73 027–73 050, 2020, conference Name: IEEE Access.

[3] S. Carpin, S. Liu, J. Falco, and K. V. Wyk, "Multi-fingered robotic grasping: A primer," *CoRR*, vol. abs/1607.06620, 2016. [Online]. Available: http://arxiv.org/abs/1607.06620

[4] C. Melchiorri, "Slip detection and control using tactile and force sensors," *IEEE/ASME Transactions on Mechatronics*, vol. 5, no. 3, pp. 235–243, Sep. 2000.

[5] K. Van Wyk and J. Falco, "Calibration and Analysis of Tactile Sensors as Slip Detectors," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 2018, pp. 2744–2751, iSSN: 2577-087X.

[6] I. Agriomallos, S. Doltsinis, I. Mitsioni, and Z. Doulgeri, "Slippage Detection Generalizing to Grasping of Unknown Objects Using Machine Learning With Novel Features," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 942–948, Apr. 2018, conference Name: IEEE Robotics and Automation Letters.

[7] M. Meier, G. Walck, R. Haschke, and H. J. Ritter, "Distinguishing sliding from slipping during object pushing," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2016, pp. 5579–5584, iSSN: 2153-0866.

[8] L. L. Salisbury and A. B. Colman, "A mechanical hand with automatic proportional control of prehension," *Medical and biological engineering*, vol. 5, no. 5, pp. 505–511, Sep. 1967. [Online]. Available: https://doi.org/10.1007/BF02479145

[9] R. S. Johansson and J. R. Flanagan, "Coding and use of tactile signals from the fingertips in object manipulation tasks," *Nat Rev Neurosci*, vol. 10, no. 5, pp. 345–359, May 2009, number: 5 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/nrn2621

[10] W. Chen, H. Khamis, I. Birznieks, N. F. Lepora, and S. J. Redmond, "Tactile Sensors for Friction Estimation and Incipient Slip Detection—Toward Dexterous Robotic Manipulation: A Review," *IEEE Sensors Journal*, vol. 18, no. 22, pp. 9049–9064, Nov. 2018, conference Name: IEEE Sensors Journal.

[11] W. A. Friedl and M. A. Roa, "Experimental Evaluation of Tactile Sensors for Compliant Robotic Hands," *Frontiers in Robotics and AI*, vol. 8, 2021.

[12] F. Veiga, H. van Hoof, J. Peters, and T. Hermans, "Stabilizing novel objects by learning to predict tactile slip," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2015, pp. 5065–5072.

[13] W. Mandil, K. Nazari, and A. Ghalamzan E, "Action conditioned tactile prediction: a case study on slip prediction," 2022. [Online]. Available: https://arxiv.org/abs/2205.09430

[14] Y. Zhang, Z. Kan, Y. A. Tse, Y. Yang, and M. Y. Wang, "Fingervision tactile sensor design and slip detection using convolutional lstm network," 2018. [Online]. Available: https://arxiv.org/abs/1810.02653

[15] G. Yan, A. Schmitz, T. P. Tomo, S. Somlor, S. Funabashi, and S. Sugano, "Detection of Slip from Vision and Touch," in *2022 International Conference on Robotics and Automation (ICRA)*, May 2022.

[16] J. Toskov, R. Newbury, M. Mukadam, D. Kulić, and A. Cosgun, "In-hand gravitational pivoting using tactile sensing," 2022. [Online]. Available: https://arxiv.org/abs/2210.05068

[17] I. Waters, D. Jones, A. Alazmani, and P. Culmer, "Utilising incipient slip for grasping automation in robot assisted surgery," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1071–1078, 2022.

[18] Z. Xia, Z. Deng, B. Fang, Y. Yang, and F. Sun, "A review on sensory perception for dexterous robotic manipulation," *International Journal of Advanced Robotic Systems*, vol. 19, no. 2, p. 17298806221095974, 2022. [Online]. Available: https://doi.org/10.1177/17298806221095974

[19] N. Wettels, A. R. Parnandi, J.-H. Moon, G. E. Loeb, and G. S. Sukhatme, "Grip Control Using Biomimetic Tactile Sensing Systems," *IEEE/ASME Transactions on Mechatronics*, vol. 14, no. 6, pp. 718–723, Dec. 2009, conference Name: IEEE/ASME Transactions on Mechatronics.

[20] G. Yan, A. Schmitz, T. P. Tomo, S. Somlor, S. Funabashi, and S. Sugano, "Detection of slip from vision and touch," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 3537–3543.

[21] K. Nazari, W. Mandil, and A. Ghalamzan E, "Proactive slip control by learned slip model and trajectory adaptation," 2022. [Online]. Available: https://arxiv.org/abs/2209.06019

[22] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," 2014. [Online]. Available: https://arxiv.org/abs/1409.1259

[23] B. Zeng, H. Liu, H. Song, Z. Zhao, S. Fan, L. Jiang, Y. Liu, Z. Yu, X. Zhu, J. Chen, and T. Zhang, "Design and slip prevention control of a multi-sensory anthropomorphic prosthetic hand," *Industrial Robot: the international journal of robotics research and application*, vol. 49, no. 2, pp. 289–300, Jan 2022.