Spiking Neural Network Based on Threshold Encoding For Texture Recognition

Haydar Al Haj Ali[1], Ali Dabbous[1], Ali Ibrahim[1,2] and Maurizio Valle[1]

***Abstract*—This paper presents a neuromorphic computing model that classifies material textures using a neural coding scheme based on threshold encoding. The proposed threshold encoding converts raw tactile data of each texture into an event-based data highlighting the spatio-temporal features needed to recognize human touch. Achieved results show that the model can categorize the input tactile signals into their corresponding material textures with high accuracy and fast inference. This work paves the way toward employing the proposed encoding method in more complex tactile based applications from the theoretical and hardware implementation aspects.**

***Index Terms*—Neuromorphic, Spatio-temporal, spiking neural network, Event-driven, Tactile perception**

I. Introduction

Human sense of touch represents a fundamental aspect in interacting with the adjoining environment on a daily basis. The human exquisite somatosensory system enables to discriminate and recognize numerous exposed bodies/actions to our skin, such as textures, objects, innocuous and noxious touches [1]. The somatosensory system involves the activation of the primary sensory tactile neurons so-called mechanoreceptors (slowly-adapting type 1: SA-1, and fast-adapting type 1: FA-1) innervating the glabrous skin of the human fingertip [2]. Consequently, the aforementioned neurons will sequentially convey detected mechanical stimuli to the somatosensory cortex by means of peripheral afferent fibers to perform information decoding [2]. Endowing robotics and prosthetic devices with such human trait, would enable performing and manipulating more complex tasks, and physically interact with other agents in an efficient way. However, mimicking the human touch artificially is still far-flung, due to many limitations and challenges [3].

Neuromorphic computing has attracted recent researches focusing on the implementation on tactile sensing systems since it features brain-inspired computing primitives [4] [5] [6] Hardware tailored neuromorphic devices exhibit an extreme computational efficiency due to their capability of performing in-memory computations where memory and computations are parallelized like neurons. This non-Von Neumann based architecture enable achieving efficient computations in terms of energy consumption and inference time. From the biological aspect, neuromorphic systems emulates in some measure the physiological behaviours of the mechanoreceptors by generating and processing spike-based encoding that are observed in natural senses [7]. Such induced encoding fetch spatio-temporal information about the applied mechanical stimuli, and that is expedient in the domain of touch since tactile perception is based on spatial and temporal perspective. Many works employed neuromorphic computing with tactile sensing systems for human touch applications such as touch modality classification [8], edge orientation classification [6], texture classification [9], [10], and object shape recognition [11]. As mentioned above, spatio-temporal features elicited from the neural encoding of the stimuli represent the center of attention in the sense of touch. For instance, textures are well known to have different micro-geometries and topology that controls their degree of smoothness/roughness. Therefore, each texture yields a distinguishable temporal feature that has a delicate interplay with its coarseness level.

Spiking neural networks (SNNs) have recently emerged as a feasible tool capable of performing synaptic learning and classification based on spatio-temporal features by means of bio-inspired learning algorithms. Neurons in SNNs interface by means of spike trains that are temporal signals in nature, therefore, SNNs are considered a potential candidate to deal with the temporal dynamics of a signal. In [9], a simple 2 fully connected layers spiking neural network has been implemented to classify 20 naturalistic textures collected from two different tactile sensors: SynTouch Biotac sensor piezo-resistive based and iCub Roboskin piezo-capacitive based under different experimental conditions. SLAYER back propagation based learning approach has been adopted to learn from the neural coding of the raw data. Authors in [10] implemented 2 layer SNN to classify 8 artificial textures collected from piezo-electric tactile sensors, using unsupervised Spike-Timing-Dependent Plasticity (STDP) as a learning approach. Experimental results report a classification accuracy more than 80%. Udaya et al. [5] proposed a 2 layer SNN endowed with homeostatic synaptic learning mechanism to classify 10 naturalistic textures under varying sensing conditions. In [12], authors employed K-nearest neighbour (KNN) to classify 10 textures. 2 features based on: coefficient of variation and average spike rate have been extracted from the neural encoding of the textures done by Izhikevic model, and used to train the classifier and perform classification. Sankar et al. [13] used a soft bio-mimetic fingertip to classify 13 artificial textures. The authors converted the raw tactile data of the textures acquired by piezoresistive sensor into neural encoding by means of

[1] Department of Electrical, Electronic, Telecommunication Engineering and Naval Architecture DITEN University of Genoa, Genova, Italy.

[2] Department of Electrical and Electronics Engineering, Lebanese International University (LIU), Beirut 1105, Lebanon

The authors acknowledge partial financial support from TACTIle feedback enriched virtual interaction through virtual realITY and beyond (Tactility) project: EU H2020, Topic ICT-25-2018-2020, RIA, ProposalID 856718.

Izhikevic model. Moreover, support vector machine (SVM) was employed to classify the textures based on hand-crafted features extracted from the encoding. Experimental results reports a classification accuracy of 98.65%.

Considerable works in the literature addressed the implementation of machine and deep learning approaches for texture classification applications. However, the deployment of such models on hardware devices encounters many challenges such as: memory requirements, computational load and energy consumption. Neuromorphic systems have been introduced to solve aforementioned problems due to their binary encoding (1: spike and 0: no spike) and dedicated hardware accelerators that supports online learning strategies based on biologically plausible learning rules. This paper presents a soft SNN implementation for texture classification based on threshold encoding. The simplicity of the proposed encoding scheme could be used in future works that target the implementation on neuromorphic or low-power edge-devices for real-time inference. The main contribution of this work can be summarized as follows:

- A neural encoding scheme based on threshold encoding that converts raw tactile information into event-representations with a salient biologically plausible features i.e. spatio-temporal information.
- A soft SNN framework based on a simple 2-layer spiking network that can be deployed on a neuromirphic device.
- The proposed system overcomes similar state of the art solution in [9] by achieving a classification accuracy of 100%.

II. Experimental Setup

In this work, an online available dataset [14] has been used for analyzing tactile sensing perception. The dataset consists of 20 naturalistic textures listed as follows: Cotton, Bath Towel, Carpet Net, Leather Fake, Wood Hard, Foam, Metal, Fiber Board, Cork, Eva, Soft Material 1 & 2, Sponge Soft, Felt, Polypropylene Thin, Polypropileno Smooth, Paper 1 & 2 and Styrofoam. A SynTouch BioTac sensor [15] piezoresistive based composed of 19 channels is used to acquire pressure readings of the textures. The BioTac sensor is attached as a passive-end effector to the KUKA LBR iiwa robot in order to acquire data under a completely controlled environment in terms of applied force and velocity. The data has been collected by placing the textures on a flat surface, meanwhile the robotic arm including the sensor performs passive sliding on the top surface of the material for 8 seconds with a constant speed of 2.5 cm/s and a sampling rate of 100 Hz [14]. The work in [9] used as well the dataset collected in [14].

III. Methodology

This section marks out a discussion with respect to the proposed neural encoding mechanism based on threshold encoding, and the architecture of the system.

A. Threshold Encoding

Tactile perception and cognitive task entails primarily the collaboration of front-end sensors, that encodes the temporal dynamic stimulus into an efficient and effective spike events. Thus emulating to certain measure the function of perceptual units of the human cognitive system [14]. Threshold encoding is found to be effective in many cases [9] [4]. It aims to encode the input stimulus using threshold crossing events in a population of sensory neurons [16]. Moreover, it mitigates the workload of the neural network in the classification part by generating a linearly separable dataset and making the inputs classifiable based on spike counts alone.

In this work, we proposed a threshold encoding mechanism that depends drastically on the distinct raw tactile value of each texture. The micro-topology of each texture collaborates in the coarseness level, thus each texture will have its distinguished temporal feature that will be used as a threshold event for the input channels.

Algorithm 1 Threshold Encoding Procedure

1. **Input**:
   - Input Channels: $N \longleftarrow 19$
   - Raw Tactile Signal of N_{th} Channel: P
2. **Output**: Event-Based Dataset D_{ev}

3.**Methodology**

for $(n = 1; n \leq M; n++)$ **do** // M: N.b of Textures
  for $(i = 1; i \leq N; i++)$ **do**
    if $P_{average}[N(i)] == P_{average}[Texture(n)]$ **then**

$$S_{N(i)}(n) = \begin{cases} 1 & \text{if } P[N(i)] \geq P_{average} \\ 0 & \text{else} \end{cases}$$

    end if
  end for
end for

4. Return $D_{ev} = \sum_{N=1}^{19} \sum_{M=1}^{20} S_{NM} \in \{0, 1\}$

Algorithm 1 presents the procedure of the preformed threshold encoding. A sequence of 20 spike trains $S(_n)$ are generated by each input channel (N) depending on the average of the raw tactile values of each texture ($P_{average} = (\sum_{i=1}^{400} P_i)/SequenceLength$). A total of 380 spike trains $S_{N(i)}(n)$ are generated, where (n) and $N(i)$ represent the number of textures and corresponding spike train source channel. Fundamentally, each spike train represents the temporal encoding of a single texture. To elaborate more, and for the sake of simplicity, S_1 (First spike train of each channel) is generated when the average of raw data corresponds to a certain type of textures (i.e. CarpetNet), whereas the rest of spike trains are affiliated to the state of different textures. The total sequence length of the encoded values is set to 400 (First 4 seconds were selected out of the total duration, with sampling rate of 100 Hz). Moreover, the values that exceed the threshold (taken to be the average of the raw data) are considered as a spike, meanwhile values below this threshold are set to be zero (non-spike). Therefore, an event-driven dataset is generated carrying salient spatio-temporal features

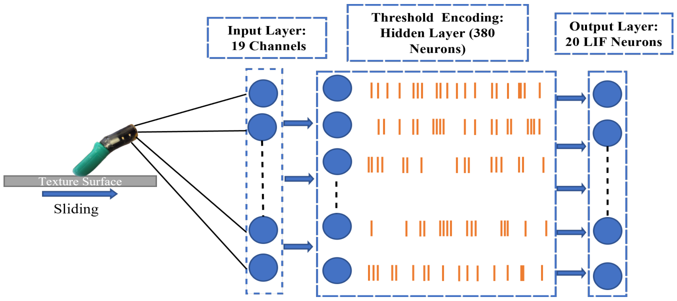

Fig. 1. An overview of the model architecture: (Left) Raw tactile data of each texture acquired from the 19 channels of the BioTac Sensor; (Middle) Threshold Encoding applied by the Hidden Layer; (Right) Decoding phase.

for each texture: population of active neuron or spike trains for each texture (spatial) and encoding that exhibit distinctive timing of spike (temporal).

B. Model Architecture

Figure 1 depicts the architecture of the system and the workflow adopted in this work. A simple 2-layer feed-forward spiking neural network comprising of a hidden and output layer has been constructed to classify the textures. The raw tactile data of each texture acquired by the 19 channels of the BioTac sensor are conveyed to the hidden layer. The hidden layer consists of 380 neurons that performs neural coding based on threshold encoding to generate event-based dataset (discussed in section III-A). Moreover, the output layer incorporates 20 LIF neurons that decodes and classifies the textures exposed to the network.

$$\tau_m \frac{du(t)}{dt} = u_{rest} - u(t) + R_m \times I(t) \tag{1}$$

Equation 1 models the voltage response of the Leaky Integrate-and-Fire neuron. Neuron membrane potential $u(t)$ generates a spike when it exceeds a defined threshold, after getting depolarized by means of the injected current $I(t)$ and the action of the membrane time constant τ_m. After emitting a spike, the voltage membrane undergo a repolarization phase until reaching the resting value V_{rest}.

The connection between hidden and output layer is derived from the selection of the active neuron population for each texture in the hidden layer according to the performed encoding mechanism. Accordingly, each texture active cluster in the hidden layer is connected to one neuron in the output layer, while the non active cluster is disregarded to reduce the complexity of the network. Therefore, each neuron in the output layer fires correspondingly to a particular texture. All the simulations and models are implemented using Brian2 simulator [17].

IV. Experimental Results

A. Textures Classification

Figure 2 shows the raster plot of the proposed network along with the firing rate of each neuron in the output layer. Apparently, the neurons in the hidden layer (figure 2: Top) are

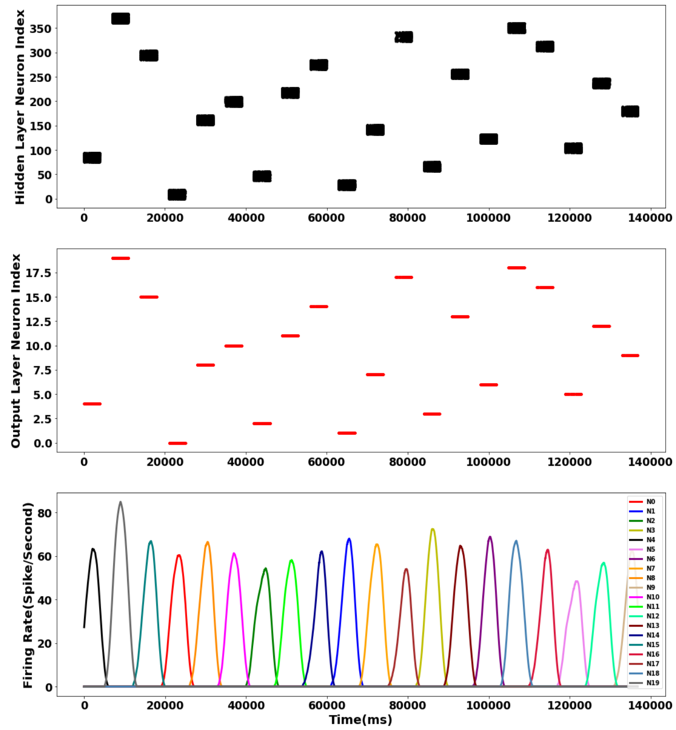

Fig. 2. Firing Activity of the network: (Top) Firing activity of the hidden layer (380 neurons). (Middle) Firing activity of the output layer (20 neurons). (Bottom) Mean firing rate of each neuron in the output layer.

sorted in an active and non active clusters per each texture, after applying the threshold encoding mechanism. The collocation process depends on the raw tactile information of each texture associated with the crossing events imposed on each generated spike train. Each neuron in the output layer fires for one texture due to the fixed-connections with the hidden layer, thus decoding successfully the received spike trains and predicting their corresponding texture. The firing rate graph (figure 2: Bottom) demonstrates a significant firing activity for each neuron in the output layer, as for every texture, only one neuron in the output layer is firing at maximum firing rate. The temporal dynamics of each texture trigger the activation of specific spike trains in the hidden layer according to the pre-defined threshold events. Consequently the variation of this temporal feature between the textures (due to typologies' assortment) collaborates in triggering disparate spike trains, thus the event-based dataset becomes linearly separable as claimed in [18]. The network is capable of recognizing all the 20 textures with an accuracy of 100%, outperforming a similar state of the art solution [9] that achieved a 94% accuracy.

B. Inference Time Response

Inference time is a metric to evaluate the overall performance and efficiency. In this work, we measured the inference time of each texture by taking the time difference (Δt) between first generated decoding spike by the output neuron t_1 and launching the touch on the surface of texture t_0 as shown in figure 3 (Top). Figure 3 (Bottom) shows a notable

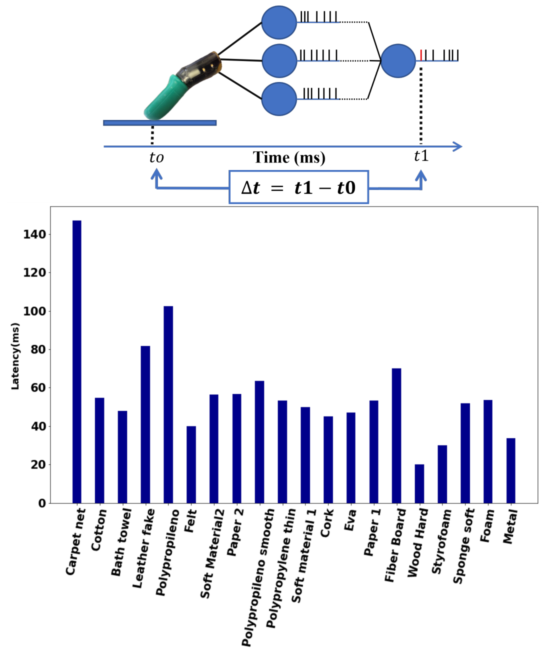

Fig. 3. Illustration of computed inference time. (Top) Output neuron inference procedure. (Bottom) Measured inference time (ms).

contrast in the inference among the textures, implying that each neuron in the output layer requires distinct duration to perform successful decoding. This duration includes the time of depolarizing the membrane potential of the neuron until emitting a spike accordingly. The computed inferences vary between 140 ms (Carpet net) and 20 ms (Wood hard) indicating a rapid response by the network compared to the state of the art solution [9]. Fundamentally, the obtained disparity is due to the variation between the micro-topology of each texture that yields a different and distinctive temporal feature for each class. The output neurons depend mainly on two factors to spike: frequency of received spikes from the pre-synaptic neuron that collaborates in the depolarization of the neuron's membrane, and the inter-spike interval (ISI) that affects the decay of the membrane potential. Moreover, a smaller inter-spike interval associated with higher spike frequency, induces a faster spike accumulation in the neuron's membrane thus a faster spike generation and a smaller inference time. Therefore, each texture with its unique topology possess a particular spike rate and ISI collaborating in different responses of their corresponding output neuron.

V. CONCLUSIONS

This paper presents a neuromorphic computing model comprised of a simple 2-layer feed forward spiking neural network to classify naturalistic textures. An efficient neural coding method based on threshold encoding has been developed to transform input raw tactile information to event-representations, and to make the network classify textures based on their spatio-temporal features. The proposed network achieved a classification accuracy of 100% overcoming a similar state of the art solution in [9] by using the event-based dataset. Future work targets the expansion and generalization of the proposed neural coding to recognize various human touch modalities.

REFERENCES

[1] V. E. Abraira and D. D. Ginty, "The sensory neurons of touch," *Neuron*, vol. 79, no. 4, pp. 618–639, 2013.

[2] J. A. Pruszynski and R. S. Johansson, "Edge-orientation processing in first-order tactile neurons," *Nature neuroscience*, vol. 17, no. 10, pp. 1404–1409, 2014.

[3] R. Dahiya, N. Yogeswaran, F. Liu, L. Manjakkal, E. Burdet, V. Hayward, and H. Jörntell, "Large-area soft e-skin: The challenges beyond sensor designs," *Proceedings of the IEEE*, vol. 107, no. 10, pp. 2016–2033, 2019.

[4] S. F. Muller-Cleve, V. Fra, L. Khacef, A. Pequeno-Zurro, D. Klepatsch, E. Forno, D. G. Ivanovich, S. Rastogi, G. Urgese, F. Zenke *et al.*, "Braille letter reading: A benchmark for spatio-temporal pattern recognition on neuromorphic hardware," *arXiv preprint arXiv:2205.15864*, 2022.

[5] U. B. Rongala, A. Mazzoni, A. Spanne, H. Jörntell, and C. M. Oddo, "Cuneate spiking neural network learning to classify naturalistic texture stimuli under varying sensing conditions," *Neural Networks*, vol. 123, pp. 273–287, 2020.

[6] A. Dabbous, M. Mastella, A. Natarajan, E. Chicca *et al.*, "Artificial bio-inspired tactile receptive fields for edge orientation classification," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2021, pp. 1–5.

[7] 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.

[8] A. Dabbous, A. Ibrahim, M. Valle, and C. Bartolozzi, "Touch modality classification using spiking neural networks and supervised-stdp learning," in *2021 28th IEEE International Conference on Electronics, Circuits, and Systems (ICECS)*. IEEE, 2021, pp. 1–4.

[9] T. Taunyazov, Y. Chua, R. Gao, H. Soh, and Y. Wu, "Fast texture classification using tactile neural coding and spiking neural network," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 9890–9895.

[10] C. Jiang, L. Yang, and Y. Zhang, "A spiking neural network with spike-timing-dependent plasticity for surface roughness analysis," *IEEE Sensors Journal*, vol. 22, no. 1, pp. 438–445, 2021.

[11] J. Kim, S.-P. Kim, J. Kim, H. Hwang, J. Kim, D. Park, and U. Jeong, "Object shape recognition using tactile sensor arrays by a spiking neural network with unsupervised learning," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2020, pp. 178–183.

[12] U. B. Rongala, A. Mazzoni, and C. M. Oddo, "Neuromorphic artificial touch for categorization of naturalistic textures," *IEEE transactions on neural networks and learning systems*, vol. 28, no. 4, pp. 819–829, 2015.

[13] S. Sankar, D. Balamurugan, A. Brown, K. Ding, X. Xu, J. H. Low, C. H. Yeow, and N. Thakor, "Texture discrimination with a soft biomimetic finger using a flexible neuromorphic tactile sensor array that provides sensory feedback," *Soft Robotics*, vol. 8, no. 5, pp. 577–587, 2021.

[14] R. Gao, T. Taunyazov, Z. Lin, and Y. Wu, "Supervised autoencoder joint learning on heterogeneous tactile sensory data: Improving material classification performance," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 10 907–10 913.

[15] J. A. Fishel and G. E. Loeb, "Sensing tactile microvibrations with the biotac—comparison with human sensitivity," in *2012 4th IEEE RAS & EMBS international conference on biomedical robotics and biomechatronics (BioRob)*. IEEE, 2012, pp. 1122–1127.

[16] J. Hopfield, "Encoding for computation: recognizing brief dynamical patterns by exploiting effects of weak rhythms on action-potential timing," *Proceedings of the National Academy of Sciences*, vol. 101, no. 16, pp. 6255–6260, 2004.

[17] M. Stimberg, R. Brette, and D. F. Goodman, "Brian 2, an intuitive and efficient neural simulator," *Elife*, vol. 8, p. e47314, 2019.

[18] Z. Pan, Y. Chua, J. Wu, M. Zhang, H. Li, and E. Ambikairajah, "An efficient and perceptually motivated auditory neural encoding and decoding algorithm for spiking neural networks," *Frontiers in neuroscience*, vol. 13, p. 1420, 2020.