Miniaturising a Neuromorphic Tactile Sensor with On-board Processing

By

Liu, Xiaohang

MSc Robotics Dissertation

School of Engineering Mathematics and Technology
UNIVERSITY OF BRISTOL
&
School of Engineering
UNIVERSITY OF THE WEST OF ENGLAND

A MSc dissertation submitted to the University of Bristol and the University of the West of England in accordance with the requirements of the degree of MASTER OF SCIENCE IN ROBOTICS in the Faculty of Engineering.

August 28, 2025

Declaration of own work

I declare that the work in this MSc dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Xiaohang Liu August 28, 2025

Ethics statement

To fill in according to the Dissertation Handbook Section 3.2.

Xiaohang Liu August 28, 2025

Miniaturising a Neuromorphic Tactile Sensor with On-board Processing

Xiaohang Liu

***Abstract*— The miniaturization of tactile sensors is extremely important for improving the flexibility of robotic hands and prosthetics. At the same time, low-power edge computing can greatly reduce the load of the robot CPU and provide real-time processing capability with lower latency. In this article, we design a smaller version (diameter 20mm) based on the standard version of NeuroTac, and evaluate its performance on edge orientation classification tasks using the SPECK neuromorphic SoC. The experiment uses an edge installed on the Dobot robotic arm to tap the central area of the sensor tip vertically downwards at different pressing depths, with a step size of 12 degrees in the edge stimulus orientation. The tactile data collected by event-based cameras is processed by a small-scale 4-layer spiking convolutional neural network to obtain angle classification results. While maintaining low power consumption (16mW) and low latency, the real-time classification accuracy of the hardware reaches 94.25%, with an average angle error of only 0.76 degrees. This proves that the tactile sensor designed in this article can accurately detect the edge stimulus orientation while reducing the overall size, providing reliable tactile information processing capabilities in the application scenarios of robotic hands and prosthetics.**

I. INTRODUCTION

Current research in robotics increasingly focuses on dexterous manipulation, where miniaturized tactile sensors and compact robotic hands play a vital role. Integrating vision-based tactile sensors[1] with biomimetic neuromorphic engineering enables efficient human–machine information transmission, fostering progress in advanced prosthetics[2]. Small onboard neuromorphic tactile sensors can not only be embedded in robotic hands of similar size to human palms, but also use event-based data to simulate the collection, processing, and transmission of tactile signals in human fingertips[3], which is of great significance for the development of dexterous hands and prosthetics.

The aims of this paper are as follows:

- Reduce the diameter of the NeuroTac sensor[4] from 30/40mm to 20mm.
- Use a spiking convolutional neural network to evaluate the performance of this miniaturized sensor in edge orientation classification tasks.

We adjust the pin layout inside the tip of NeuroTac and reduces the overall size of the tip, in conjunction with the SPECK neuromorphic SoC for the collection and processing of tactile information. When evaluating the VBTS of this small-sized neural morphology, we used an edge orientation classification task. The straight edge installed on the Dobot robotic arm strikes the contact surface of the tip vertically downwards above the sensor, with the edge stimulus direction set from 0 degrees to 168 degrees and a step size of 12 degrees. Adding position disturbance and pressing depth disturbance during the tapping process enables the sensor to perceive edge stimulus orientations beyond the center point of the tip and fixed pressing depth. The tactile data collected by the event-based camera inside the tactile sensor is input into a 4-layer SCNN, which determines the angle of the edge stimulus relative to the sensor based on the number of spikes emitted by the spiking neurons in the last layer. The final experimental results indicate that the sensor can perform real-time inference with extremely low power consumption and reliable accuracy, and is expected to be embedded in smaller robotic hands to promote further development in the fields of robot manipulation and prosthetic research.

II. RELATED WORKS

Research on tactile sensors is not only a key technology for robots to move towards humanoid intelligence[5], but also a core support for multiple fields such as medical rehabilitation[6]. To enable robots to perform various tasks as well as humans, the implementation of tactile perception is extremely important, especially by embedding neural tactile perception into bionic robotic hands[7]. At present, research has achieved the integration of artificial retina and human nervous system by combining neural morphology perception technology with visual technology[8], which also proves that neural morphology tactile perception has the potential to be applied to research in the field of prosthetics[9][10][11], helping people with disabilities restore their natural touch.

VBTS, such as TacTip[12][13], GelSight[14][15][16], currently stand out among traditional piezoelectric tactile sensors, piezoresistive tactile sensors, capacitive tactile sensors, and other tactile sensors due to their ultra-high spatial resolution. However, due to the inherent frame rate limitation of its internal camera, the delay in the time dimension is its weakness.

The emergence of event-based cameras[17][18] has filled the gap in time resolution of traditional frame-based cameras. Event cameras such as Inivation mini-eDVS, DVS240, and Prophesee GenX320 independently and asynchronously perceive brightness changes for each pixel, and reflect these changes at the pixel level through the output sparse event stream. This unique working method naturally gives it ultra-high time resolution and low latency[19], while also being more in line with biomimetic design principles. NeuroTac combines the biomimetic hardware structure of TacTip and uses an event-based camera as its camera module, which demonstrates extremely high accuracy and fast response speed in natural texture classification tasks[20]. Similarly, Evetac[21] is a tactile sensor that replaces RGB cameras with event-based cameras. With the help of the event-based

camera, it can perceive vibrations up to approximately 498 Hz in real-time, reconstruct shear forces in real-time, and achieve closed-loop anti-skid control through learning methods; Compared with frame-based schemes, it significantly reduces data rate.

Spiking neural networks[22] transmit information between neurons through discrete spikes, and are considered the third generation of artificial neural networks because they work more closely to the biological nervous system. Due to its event driven, spatiotemporal information encoding, and low-power characteristics[23], SNN is highly suitable for processing sparse dynamic signals output by tactile sensors[24]. It has shown great potential in fields such as texture recognition, edge detection, sliding detection, and neural prosthetics. In the future, it is expected to become the core technology for tactile perception in robots and neuromorphic systems. Spiking Convolutional Neural Network[25] is a type of network that combines the structure of CNN with the spatiotemporal spiking signal processing mechanism of SNN. It combines the characteristics of both neural networks and can effectively extract features in the spatiotemporal domain[26].

The most common neuron model in SNN is the integrate-and-fire (IAF) neuron, which accumulates input current into a membrane potential and emits a spike when the threshold is reached, after which the potential resets[27]. While this mechanism enables efficient temporal encoding, it also introduces a challenge for training, as the spiking function is non-differentiable. To address this, surrogate gradient methods have been proposed[28], in which the discontinuous spike function is replaced by a differentiable approximation during backpropagation. Common surrogate functions include sigmoid and triangular forms, but studies have shown that periodic exponential approximations often provide faster convergence and higher accuracy in temporal classification tasks.

However, studies have shown that deploying SNN on traditional von Neumann architecture GPUs is limited by hardware computing methods[29], which cannot fully exploit the sparsity and event driven advantages of SNN. Only by deploying it on chips with neural architecture, such as Loihi[30] manufactured by Intel and SPECK[31] manufactured by SynSense, can it significantly outperform traditional ANN in terms of energy consumption and latency. The fundamental reason for this is that the neural architecture integrates computation and storage, using event driven spike-based communication, and does not require frequent data transfer during the computation process. George et al. have deployed SNN on the Loihi2 chip and achieved human level texture recognition capability in natural texture classification tasks with NeuroTac in just 84ms, and more importantly, achieved extremely low power consumption[20].

There have been many studies on tactile sensors for edge orientation classification tasks. Aquilina et al. used a TacTip sensor to systematically collect edge stimulus direction and radial position, and achieved a performance with an average angle error of only 0.98 degrees through a KNN[32]. However, the sensors are equipped with traditional frame-based cameras, which leaves room for improvement in terms of energy consumption and latency. Macdonald et al. used the NeuroTac[33], a neuromorphic tactile sensor, and SNN and KNN based on unsupervised STDP to complete edge orientation classification tasks for pressing and sliding. Compared to KNN directly processing the raw data collected by tactile sensors, the feature extraction through three layers of SNN improved the accuracy of edge orientation classification for press based methods from 54% to 62.5%. However, there is still a lot of room for improvement in this accuracy, and the tap position and depth in the experiment should not be limited to specific situations. In addition, the diameter of this version of the NeuroTac is 30mm, and to embed it in the dexterous hands of robots it needs to be further reduced in size.

III. RESEARCH METHODOLOGY

A. Miniaturized Sensor Design

Our neuromorphic VBTS follows the design concept of the standard version of NeuroTac, but with some differences, mainly composed of tips, LEDs, and event-based cameras, as shown on the left side of Fig. 1. The tip is a black circular rubber dome with many protruding white pins inside. When the flexible contact surface of the sensor deforms due to contact with an object, these white pins will correspondingly produce displacement in different directions. This is an important process for the sensor to convert tactile information such as the position, magnitude, and direction of the contact force into visual information that can be captured by the camera. In the design of the tip, as shown on the right side of Fig. 1, we have made the following important adjustments:

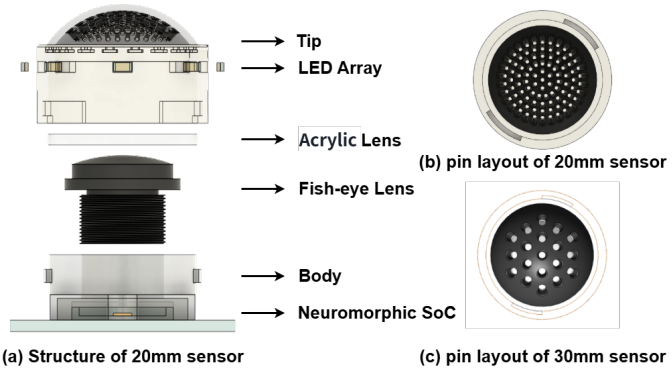

Fig. 1. **(a)** Structural diagram of a 20mm diameter neuromorphic tactile sensor. The sensor consists of a flexible tip with internal pins (filled with transparent elastic silicone), a fisheye wide-angle macro lens, and a CCD module directly connected to a neuromorphic architecture chip. **(b)** The tip after adjusting the number and arrangement of internal pins. **(c)** The internal pin layout of the 30mm diameter NeuroTac tip.

- Reduced the overall size. The diameter has been reduced from two versions of 40mm and 30mm to 20mm.
- The height of the dome has been reduced, making it shallower. This means that when the tip contact surface comes into contact with external objects and deforms,

more pins inside will be displaced, providing richer tactile information. At the same time, this will also reduce the overall thickness of the sensor, allowing it to adapt to smaller robotic hands.

- Increased the density of pins near the center point and added some relatively chaotic factors to the pin arrangement layout. The regular version of NeuroTac increases the number of internal pins by 6 per circle, giving the overall appearance of a staggered arrangement. We have increased the number of pins in the third circle from 12 to 15, which maximizes the density of pins near the center point of the tip contact surface and breaks the pattern of staggered pin arrangement. The purpose of this adjustment is also to generate richer and more complex tactile information when the contact surface deforms.

The function of LEDs is to evenly illuminate the inside of the tip, enhance the contrast between the white pins and the black inner surface, and thus make the movement of the white pins more prominent. During the process of adjusting the LEDs, we moved their position from below the acrylic glass plate (on the same side as the camera) to the side. This not only further compresses the internal space of the sensor, but more importantly, the modification of the optical path reduces the light reflection effect of the acrylic glass plate, making the data collected by the event camera clearer.

The camera module still uses event-based cameras, which maintain the advantages of high dynamic range, ultra-high time resolution, low latency, and low power consumption in the process of collecting the displacement of pins on the inner surface of the tip. Differently, the event-based camera "SPECK Module" used here integrates a neuromorphic processing chip, with a CCD embedded inside the chip and directly connected to it. Compared to standalone event-based cameras such as Inivation mini-eDVS and DAVIS240, this can reduce data transmission latency[31], and more importantly, further compress the size of sensors including data processing modules, making it possible to transmit processed advanced tactile information downstream using a single FPC line.

B. Spiking Convolutional Neural Network Architecture

We implemented a 4-layer Spiking Convolutional Neural Network, as shown in Fig. 2, on the SPECK chip to process tactile events in Address Event Representation (AER) format. The model combines conventional CNN feature extraction with spiking neuron dynamics for low-power spatiotemporal processing. The input tensor has size [2, 128, 128], representing positive and negative polarity events at 128×128 resolution. To meet the chip's storage limit (maximum output size 64), the first convolutional stage downsamples to 64×64.

Each stage consists of a convolutional layer, an IAF spiking neuron layer, and a pooling layer, as shown in Table I. The IAF neurons convert continuous activations into sparse spike sequences. Surrogate gradients (Periodic Exponential) are used to enable backpropagation training. The final stage flattens the feature maps and connects to a fully connected layer with 15 output neurons, corresponding to the 15 edge orientation classes.

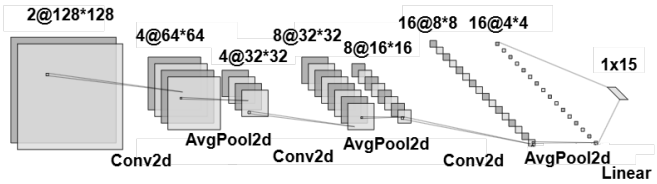

Fig. 2. Structure diagram of SCNN. The raw data captured by the event-based camera is flattened into a 1D vector after three convolutional layers, spiking neuron layers, and pooling layers, and finally output spikes are generated by the corresponding spiking neurons.

TABLE I

Stage	Layer Type	Output Size	Notes
Input	Event tensor	[2, 128, 128]	Positive/negative polarity channels
1	Conv2d → IAF → AvgPool	[4, 32, 32]	Downsample to fit chip memory limit
2	Conv2d → IAF → AvgPool	[8, 16, 16]	Feature extraction
3	Conv2d → IAF → AvgPool	[16, 4, 4]	Spatiotemporal encoding
4	Flatten → Linear → IAF	[15]	Output spiking classification

C. Experiment Setup & Dataset

To obtain the dataset, we installed an edge stimulus onto the end of the Dobot MG400 robotic arm and fixed the sensor to the bracket on the desktop, as shown in Fig. 3. The Dobot MG400 is a small desktop robotic arm with 4-DOF, achieving a repeat positioning accuracy of 0.05mm. This makes it very suitable as a data collection platform for the miniaturized tactile sensor in this article, minimizing the noise caused by the positioning error of the robotic arm. We use the CRI Python library in the developed at Bristol Robotics Laboratory to control the movement of the Dobot robotic arm.

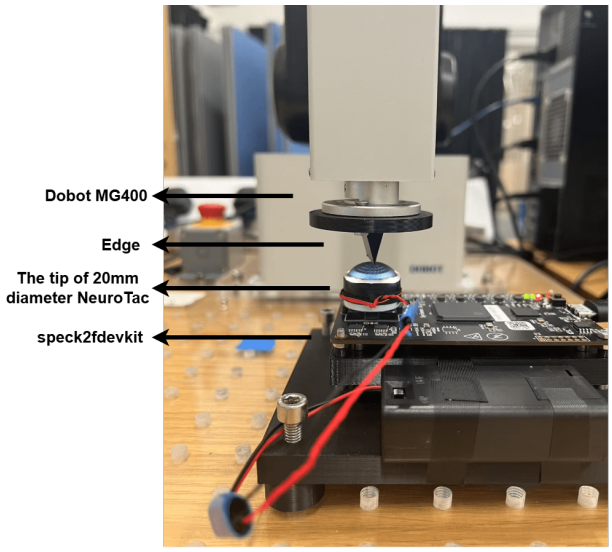

Fig. 3. A 20mm diameter tactile sensor tip installed on a neuromorphic SoC (speck2fdevkit). During the data collection process, the edge installed at the end of the Dobot MG400 was vertically tapped downwards at different positions on the sensor tip with different pressing depths.

The data collection work is completed by the interface in the Samna library provided by SynSense company. The event stream (pixel level brightness changes) collected by the event-based camera is visualized based on Samnagui, and an event stream transmission channel is established with the PC through the TCP port. After the event routing is completed, the event stream is transmitted to the data collection node, which is continuously monitored and collected by the PC.

The edge orientation classification task in this article focuses on the impact action of the edge stimulus on the sensor tip. The starting position of the edge is located about 1mm directly above the central area of the sensor, and the tip contact surface is pressed vertically downwards at different horizontal angles at a constant speed (approximately 2mm/s). The horizontal angle ranges from 0 degrees to 168 degrees, with a step size of 12 degrees. At the same horizontal angle, there will be a left-right deviation of the edge stimulus as a positional disturbance, with a distance of ± 3mm from the center point of the tip and a step size of 0.1mm. We also add disturbance to the pressing depth of the robotic arm, with a pressing depth ranging from 2mm to 3mm and a step size of 0.1mm. The collection of tactile data starts from the downward movement of the edge and continues to collect data for the next 3000ms, including the complete event flow generated during one edge pressing and lifting processes. 15*61*11=10065 pieces of data were collected, with data from each angle divided into training, validation, and testing sets in a ratio of 7:1.5:1.5.

D. Offline Training & QAT Tune

The offline training and hardware deployment of the SCNN in this article use the open-source libraries Sinabs and Samna provided by SynSense. Sinabs is an open-source PyTorch framework library used for building, training, and simulating SCNN, and then combined with Samna to generate hardware configuration files for speck2fdevkit.

Sinabs provides two methods for training SCNN: the first is to convert the trained CNN into SCNN. This method is faster in training and can use mature ANN architectures and optimization techniques because it does not require simulating temporal spikes. However, the resulting SCNN is sensitive to neuron parameters, and inappropriate threshold voltages or attenuation constants can significantly reduce classification accuracy. The second approach is to directly train SCNN through Backward Propagation through Time[34]. This training method incorporates spiking neurons when designing the model and solves the problem of non differentiable spike firing through the surrogate gradient method. This approach can minimize the abnormal interruption of signals in the inference process of neural networks due to insufficient spike quantity and intensity. Therefore, this article adopts the second training method, as shown in the SCNN architecture in III-B.

During training, we extract the pressing stage of each data in the dataset as input to the neural network. At the same time, in order to adapt to the batch tensor input of convolutional layers, we use a 10ms time window to accumulate the event stream in the input data into frames. The predicted edge stimulus orientation category is given by the spiking neuron with the highest number of spikes issued in the last spiking neuron layer of SCNN within the duration of the sample data. In addition, 5 PATIENCE early stopping strategies were added during the training process to prevent overfitting of the model during training.

When SCNN is trained on GPU, the weights of the model and the membrane potentials of spiking neurons are stored in float32 precision values. But when SCNN is deployed on speck2fdevkit, the weights of the model will be quantized to 8-bit integers, and internal variables such as neuron membrane potential will be quantized to 16 bits. If the trained SCNN is not subjected to Quantitative-Aware Training, the accuracy of the model will suffer significant losses after deployment on hardware. Here, we train the model with FP32 accuracy and insert pseudo quantization nodes into the SCNN model to obtain the QAT model. We then fine tune the QAT model on the dataset. The pseudo quantization node simulates the quantization round operation, treating this error as training noise. During QAT fine-tuning, the model adapts to this noise, thereby reducing the loss of accuracy when the model weights are quantized to 8-bit integers during deployment. It should be noted that periodic calibration is required during the QAT process. If the quantization scale is only initially calibrated once, the original quantization interval may become inappropriate as the weight or threshold range changes.

E. Real-time Classification

After the trained SCNN is deployed on speck2fdevkit, with the help of a neuromorphic hardware architecture designed specifically for asynchronous event input, tactile information can be extracted from the raw event stream directly from DVS or from the raw event stream input from PC to obtain predicted values for edge orientation classification without the need to convert the raw event stream into a sequence of spiking frames.

When deploying SCNN to speck2fdevkit, we can choose to add a readout layer afterwards. This layer has 15 channels, and counts the number of events input to each channel's neurons in each clock cycle. The predicted value for edge orientation classification is generated by searching for the index of the channel with the highest cumulative spike count within a clock cycle. The addition of the readout layer allows us to adjust the output frequency of hardware classification results by customizing the clock frequency, thereby achieving the advantage of low latency.

IV. RESULTS

A. Inspection of Data

When the edge stimulus hits the tip contact surface vertically, each event in the raw event stream captured by the event-based camera contains information such as the coordinates (x and y), polarity (brightening or darkening), and timestamp of the pixel where brightness changes occurred. In this experiment, we only used data collected during

edge pressing (data within the range of 300-800ms in the complete 3000ms data). Fig. 4 shows a line graph of four randomly selected event streams using a 10ms time window to accumulate the number of spikes. As expected, under external stimuli from different angles in the edge stimulus direction, different numbers of pins inside the sensor tip move, producing different spike firing rates and exhibiting recognizable features in the time dimension.

In addition, the original event stream can also provide additional features in the spatial dimension. The different deformation modes of the sensor tip can cause the internal pins to produce different directional motion patterns. Fig. 5 shows the distribution pattern of the original event flow within the sensor area. We can clearly see that the internal pins of the tip move towards both sides along the indentation on the contact surface. Under different directions of edge stimulation, the overall movement trend of the pins varies.

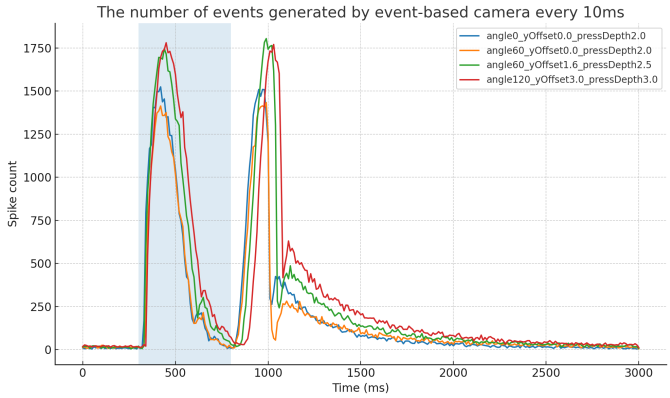

Fig. 4. The four raw event streams collected by the event-based camera accumulate the number of events generated within a 10ms time window. The blue area represents the data during the period of 300ms-800ms, corresponding to the process of vertically tapping the sensor tip contact surface with the edge stimulus.

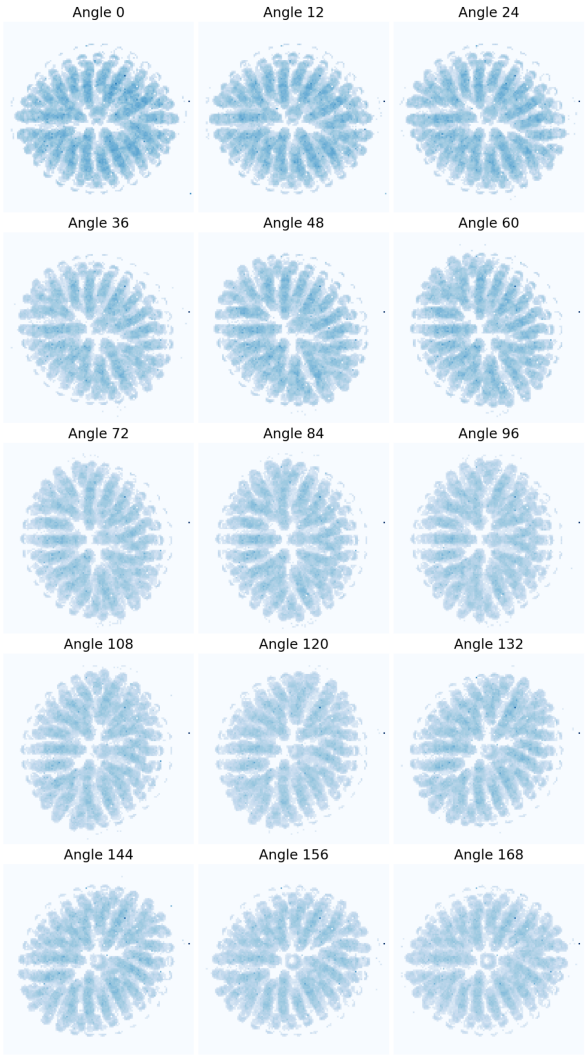

Fig. 5. The raw event stream data collected by the event-based camera accumulates events within 300ms 800ms in a 128 * 128 resolution space based on pixel coordinates x and y.

B. Edge Orientation Classification Performance

Fig. 6 shows the changes in loss and accuracy of the model during the training process. The blue curve represents the loss of training set, the orange curve represents the loss of validation set, and the green dashed line represents the accuracy of validation set. It can be observed that as the number of training epochs increases, the training set loss continues to decrease and stabilize, while the validation set loss also maintains an overall decrease before the 18th epoch, and then begins to rebound. The accuracy of the validation set gradually improved and reached its peak in the 18th epoch. This indicates that the model gradually learns effective features during the training process and exhibits good generalization ability on the validation set. Due to the implementation of an early stop strategy with 5 PATIENCEs, the training ended in the 23rd epoch.

The accuracy of the miniaturized validation set designed in this article gradually improves, and the classification accuracy of the neuromorphic tactile sensor combined with SCNN on the GPU test set reaches 99.67%. However, due to the accuracy quantization loss of weight and other parameters when deploying the model on speck2fdevkit for real-time inference, the classification accuracy drops to 91.63%. When using the QAT fine-tuned model for onboard real-time inference, the edge direction classification accuracy can still maintain 94.25%, as shown in Fig. 7, with an average angle error of no more than 0.76 degrees. From Fig. 7, it can be clearly seen that the SCNN trained on the GPU has the smallest accuracy loss on speck2fdevkit after 4 epochs of QAT fine-tuning.

We can clearly see from the confusion matrix (Fig. 8) that most of the edge stimulus directions can be accurately identified, and the misjudgment results of direction prediction are often near the true category. However, it should be noted that the sensor has poor ability to recognize edge stimulus at 48 degrees.

Besides, here we further examine what the SCNN has learned by visualizing both the learned convolutional kernels

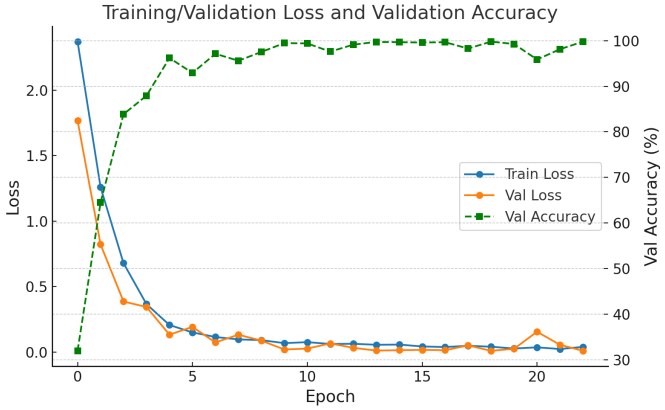

Fig. 6. The learning curve of SCNN model during simulation training on GPU, including training loss, validation loss, and validation accuracy.

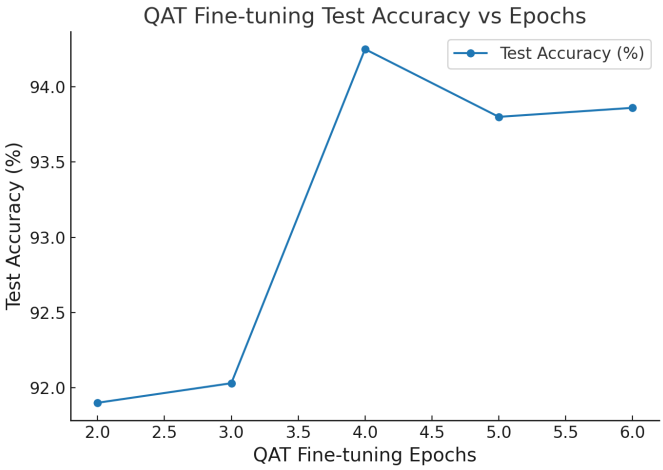

Fig. 7. The relationship between hardware real-time inference accuracy and epochs during QAT fine-tuning of SCNN models simulated and trained on GPU.

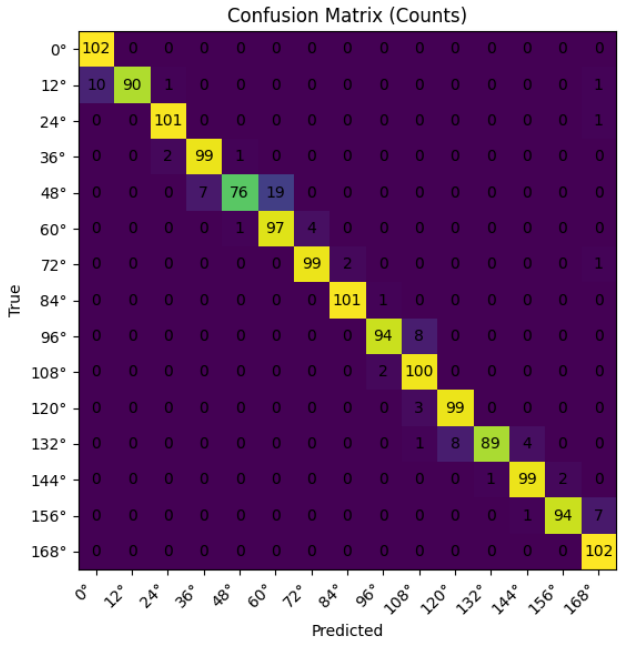

Fig. 8. The confusion matrix obtained when deploying the SCNN fine tuned by QAT after 4 epochs for on-board real-time classification of edge stimulus orientations.

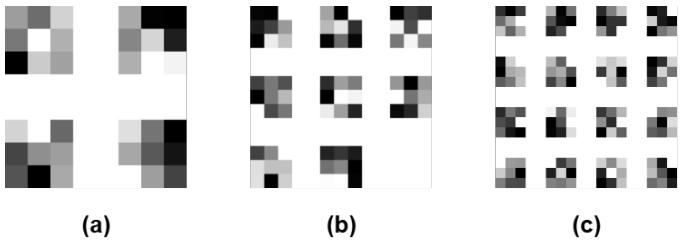

Fig. 9. **Learned convolutional kernels across layers.** (a) first layer (top-4 merged kernels), (b) second layer (top-8), (c) third layer (top-16). The kernels exhibit clear orientation selectivity and polarity contrast, indicating that the SCNN learns edge-direction filters aligned with the task.

and the intermediate feature maps. As shown in Fig. 9, the first-layer kernels display clear orientation selectivity with polarity-opponent structure (bright–dark vs. dark–bright), resembling Gabor-like edge detectors that are well matched to an edge-orientation task. Deeper layers become more compact and abstract, indicating a progressive integration of low-level edges into higher-level orientation templates. Together, these kernels suggest that the network is encoding direction-specific evidence directly from the dual-polarity event input.

To link kernels to actual network behavior, Fig. 10 shows intermediate feature maps for a representative test sample at the time bin with the strongest response. In the first layer, activations emphasize local edge-like patterns around the indentation produced by the edge contact. The second layer aggregates these responses into oriented bands and repeated motifs, which align with the stimulus direction. By the third layer, activations become sparser and more selective, concentrating on regions most informative for the final class decision. For readability, we visualize only the top-4/8/8 channels (by activation norm) from conv1/conv2/conv3; the full tensors contain more channels.

C. Power Usage

The peak power consumption of the SCNN model without QAT fine-tuning when running the test set on GPU reached 22.2W, with an average power consumption of up to 20.2W. After QAT fine-tuning, the peak power consumption of SCNN deployed on the micro neuromorphic tactile sensor designed in this paper for real-time inference was only 16mW, as shown in Fig. 11, which is only 0.07% of GPU energy consumption.

V. DISCUSSION AND FUTURE WORK

Our miniaturized neuromorphic VBTS with on-chip SCNN achieved 94.25% edge orientation classification accuracy and an average error of 0.76°, a significant improvement over the standard NeuroTac (62.5% / 4.44°). This enhancement is mainly attributed to two factors: (i) irregular pin

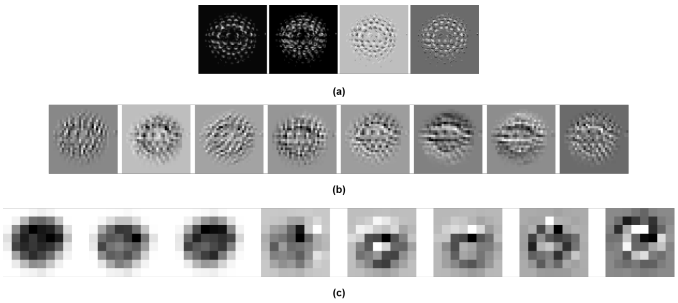

Fig. 10. **Intermediate feature maps for a representative test sample.** (a) conv1: top-4 activations, (b) conv2: top-8, (c) conv3: top-8. Visualized at the time bin with the strongest response for this sample.

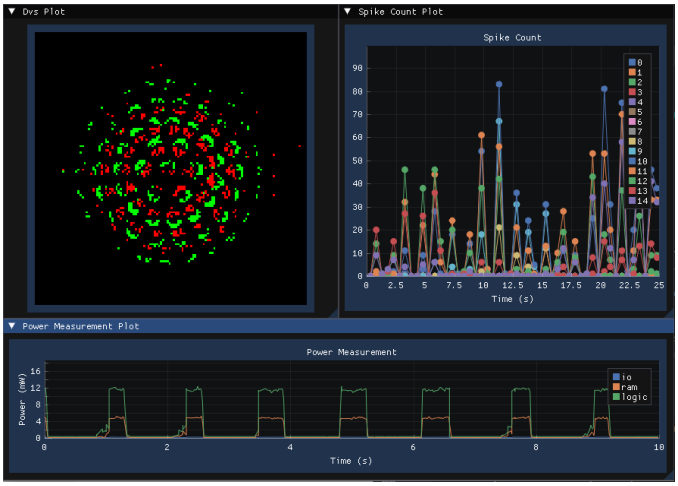

Fig. 11. The real-time monitoring interface for neuromorphic hardware provided by Samnagui, including a DVS input visualization window (top left), a spike issuance count window for the last layer of spiking neurons (top right), and a hardware energy consumption window (bottom).

arrangement at the sensor tip, which increases the distinctiveness of tactile patterns, and (ii) the SCNN architecture, which captures both spatial and temporal features in a biologically inspired manner.

Beyond numerical performance, these results have broader implications. For prosthetic applications, the miniaturized and low-power sensor can provide users with more precise tactile feedback without increasing weight or energy burden, supporting long-term daily use. For robotic manipulation, the ability to perform real-time tactile processing on a neuromorphic chip enables robots to operate with greater dexterity in unstructured environments, even without high-power GPUs. More generally, the extremely low power consumption (0.07% of GPU) highlights the promise of neuromorphic tactile sensing for scalable and energy-efficient artificial tactile systems.

From the confusion matrix in Fig. 8, it can be seen that the tactile sensor in this article has relatively poor classification performance for edge stimulus directions at 48 degrees, and some data is misclassified as 60 degrees. We believe that the reason for this phenomenon is that the center point of the edge contact object and the center point of the tactile sensor tip were not fully calibrated during data collection, resulting in a certain similarity between the data in the 48 degree edge stimulus direction and the data in the 60 degree direction.

Due to the fact that the neuromorphic chip speck2fdevkit used in this article is a CQFP packaged version, the CCD is fixed on a larger PCB, which greatly limits the miniaturization of tactile sensors to a certain extent. In future work, we will use small-sized wide-angle lenses with smaller diameters and thinner thicknesses, and use Camera Module packaged versions of SPECK or smaller event-based camera GenX320, which can help further reduce the internal space of the sensor and increase its flexibility. In addition, the edge orientation classification ability of miniaturized tactile sensors under sliding contact conditions should be explored, as well as how low the delay can be when the sensor extracts tactile information in real-time. These will have greater significance for applying tactile sensors with onboard processing capabilities to the fields of dexterous hands and prosthetics in robots.

VI. CONCLUSIONS

We propose a 20mm diameter neuromorphic tactile sensor with onboard processing capabilities, and validate its excellent performance with the help of SCNN through edge orientation classification tasks. When simulating edge orientation classification tasks on GPUs, the network achieved an accuracy of 99.67%. Even when deployed on hardware for real-time edge orientation classification, the classification accuracy remained at 94.65% after QAT fine-tuning, with an average angle error of less than 0.76 degrees, and the energy consumption for completing one inference is only 16mW. This will contribute to the development of more biomimetic robotic dexterous hands and research progress in the field of prosthetics in the future.

REFERENCES

[1] S. Zhang, Z. Chen, Y. Gao, W. Wan, J. Shan, H. Xue, F. Sun, Y. Yang, and B. Fang, "Hardware technology of vision-based tactile sensor: A review," *IEEE Sensors Journal*, vol. 22, no. 22, pp. 21 410–21 427, 2022.

[2] S. Li, Z. Wang, C. Wu, X. Li, S. Luo, B. Fang, F. Sun, X.-P. Zhang, and W. Ding, "When vision meets touch: A contemporary review for visuotactile sensors from the signal processing perspective," *IEEE Journal of Selected Topics in Signal Processing*, vol. 18, no. 3, pp. 267–287, 2024.

[3] T. Liu, G. Brayshaw, A. Li, X. Xu, and B. Ward-Cherrier, "Neuromorphic touch for robotics-a review," *Neuromorphic Computing and Engineering*, 2025.

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

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

[6] L. Osborn, W. W. Lee, R. Kaliki, and N. Thakor, "Tactile feedback in upper limb prosthetic devices using flexible textile force sensors," in *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*. IEEE, 2014, pp. 114–119.

[7] Z. Kappassov, J.-A. Corrales, and V. Perdereau, "Tactile sensing in dexterous robot hands," *Robotics and Autonomous Systems*, vol. 74, pp. 195–220, 2015.

[8] K. A. Zaghloul and K. Boahen, "A silicon retina that reproduces signals in the optic nerve," *Journal of neural engineering*, vol. 3, no. 4, p. 257, 2006.

[9] Y. Wu, Y. Liu, Y. Zhou, Q. Man, C. Hu, W. Asghar, F. Li, Z. Yu, J. Shang, G. Liu, *et al.*, "A skin-inspired tactile sensor for smart prosthetics," *Science Robotics*, vol. 3, no. 22, p. eaat0429, 2018.

[10] I. Dhar, B. B. Choudhury, B. Sahoo, and S. K. Sahoo, "A comprehensive review on advances in sensor technologies for prosthetic palms," *Spectrum of Engineering and Management Sciences*, vol. 3, no. 1, pp. 253–261, 2025.

[11] O. Diaz-Hernandez, "A worldwide research overview of artificial proprioception in prosthetics," *PLOS Digital Health*, vol. 4, no. 4, p. e0000809, 2025.

[12] B. Ward-Cherrier, N. Pestell, L. Cramphorn, B. Winstone, M. E. Giannaccini, J. Rossiter, and N. F. Lepora, "The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies," *Soft robotics*, vol. 5, no. 2, pp. 216–227, 2018.

[13] N. F. Lepora, "Soft biomimetic optical tactile sensing with the tactip: A review," *IEEE Sensors Journal*, vol. 21, no. 19, pp. 21 131–21 143, 2021.

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

[15] A. C. Abad and A. Ranasinghe, "Visuotactile sensors with emphasis on gelsight sensor: A review," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 7628–7638, 2020.

[16] I. H. Taylor, S. Dong, and A. Rodriguez, "Gelslim 3.0: High-resolution measurement of shape, force and slip in a compact tactile-sensing finger," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 10 781–10 787.

[17] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis, *et al.*, "Event-based vision: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 1, pp. 154–180, 2020.

[18] B. Chakravarthi, A. A. Verma, K. Daniilidis, C. Fermuller, and Y. Yang, "Recent event camera innovations: A survey," in *European Conference on Computer Vision*. Springer, 2024, pp. 342–376.

[19] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128$\times$ 128 120 db 15 μs latency asynchronous temporal contrast vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.

[20] G. Brayshaw, B. Ward-Cherrier, and M. J. Pearson, "A neuromorphic system for the real-time classification of natural textures," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 1070–1076.

[21] N. Funk, E. Helmut, G. Chalvatzaki, R. Calandra, and J. Peters, "Evetac: An event-based optical tactile sensor for robotic manipulation," *IEEE Transactions on Robotics*, 2024.

[22] J. D. Nunes, M. Carvalho, D. Carneiro, and J. S. Cardoso, "Spiking neural networks: A survey," *IEEE access*, vol. 10, pp. 60 738–60 764, 2022.

[23] K. Yamazaki, V.-K. Vo-Ho, D. Bulsara, and N. Le, "Spiking neural networks and their applications: A review," *Brain sciences*, vol. 12, no. 7, p. 863, 2022.

[24] 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.

[25] C. Lee, G. Srinivasan, P. Panda, and K. Roy, "Deep spiking convolutional neural network trained with unsupervised spike-timing-dependent plasticity," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 11, no. 3, pp. 384–394, 2018.

[26] Y. Cao, Y. Chen, and D. Khosla, "Spiking deep convolutional neural networks for energy-efficient object recognition," *International Journal of Computer Vision*, vol. 113, no. 1, pp. 54–66, 2015.

[27] A. N. Burkitt, "A review of the integrate-and-fire neuron model: I. homogeneous synaptic input," *Biological cybernetics*, vol. 95, no. 1, pp. 1–19, 2006.

[28] E. O. Neftci, H. Mostafa, and F. Zenke, "Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks," *IEEE Signal Processing Magazine*, vol. 36, no. 6, pp. 51–63, 2019.

[29] A. Bhattacharjee, R. Yin, A. Moitra, and P. Panda, "Are snns truly energy-efficient?—a hardware perspective," in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2024, pp. 13 311–13 315.

[30] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," *Ieee Micro*, vol. 38, no. 1, pp. 82–99, 2018.

[31] O. Richter, Y. Xing, M. De Marchi, C. Nielsen, M. Katsimpris, R. Cattaneo, Y. Ren, Y. Hu, Q. Liu, S. Sheik, *et al.*, "Speck: A smart event-based vision sensor with a low latency 327k neuron convolutional neuronal network processing pipeline," *arXiv preprint arXiv:2304.06793*, 2023.

[32] K. Aquilina, D. A. Barton, and N. F. Lepora, "Principal components of touch," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 4071–4078.

[33] F. L. Macdonald, N. F. Lepora, J. Conradt, and B. Ward-Cherrier, "Neuromorphic tactile edge orientation classification in an unsupervised spiking neural network," *Sensors*, vol. 22, no. 18, p. 6998, 2022.

[34] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 2002.