NUS National University of Singapore

School of Computing

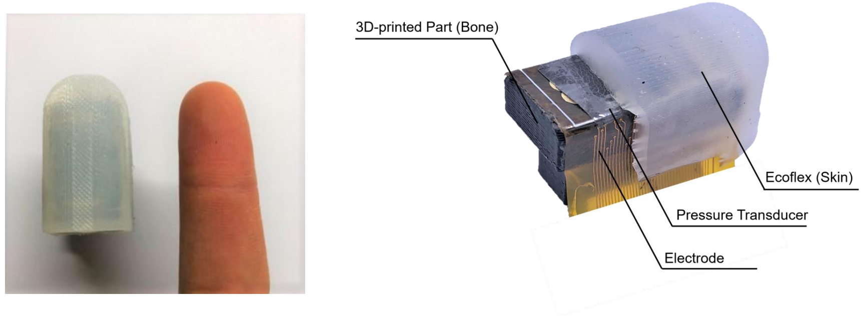
ROBOTICS SCIENCE AND SYSTEMS

Event-Driven Visual-Tactile Sensing and Learning for Robots

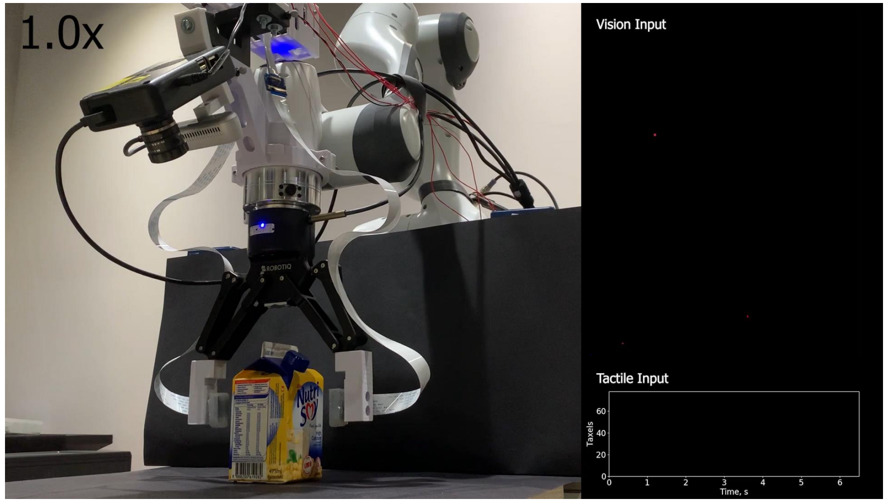
Tasbolat Taunyazov, Weicong Sng, Hian Hian See, Brian Lim, Jethro Kuan, Abdul Fatir Ansari, Benjamin C.K. Tee, Harold Soh

Event-Driven Visual-Tactile Sensing and Learning for Robots

1



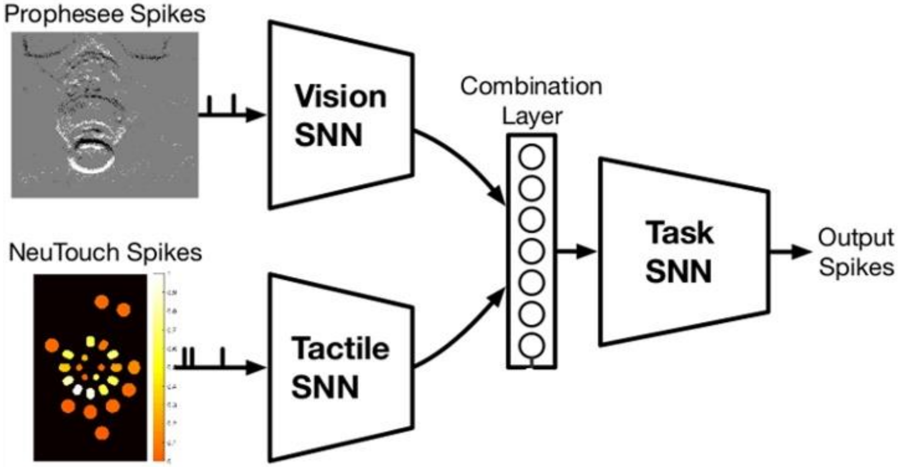
Event-based tactile sensor with graphene-based piezoresistive transducer



VT-SNN is tested for two different robotic tasks:

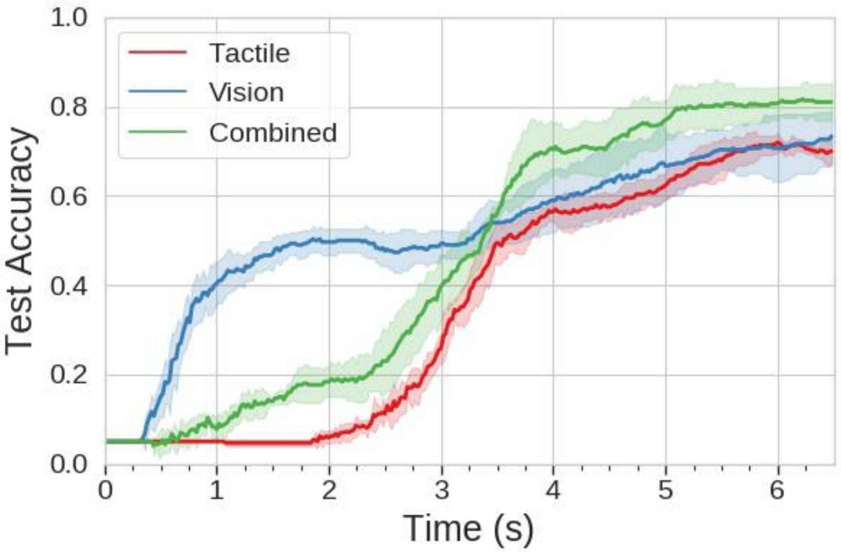
- Container and weight classification
- Rotational slip detection

2



Visual-Tactile Spiking Neural Network

$$L_\omega = \frac{1}{2}\sum_{n=0}^{N_o}\left(\sum_{t=0}^{T}\omega(t)s^n(t) - \sum_{t=0}^{T}\omega(t)\tilde{s}^n(t)\right)^2$$



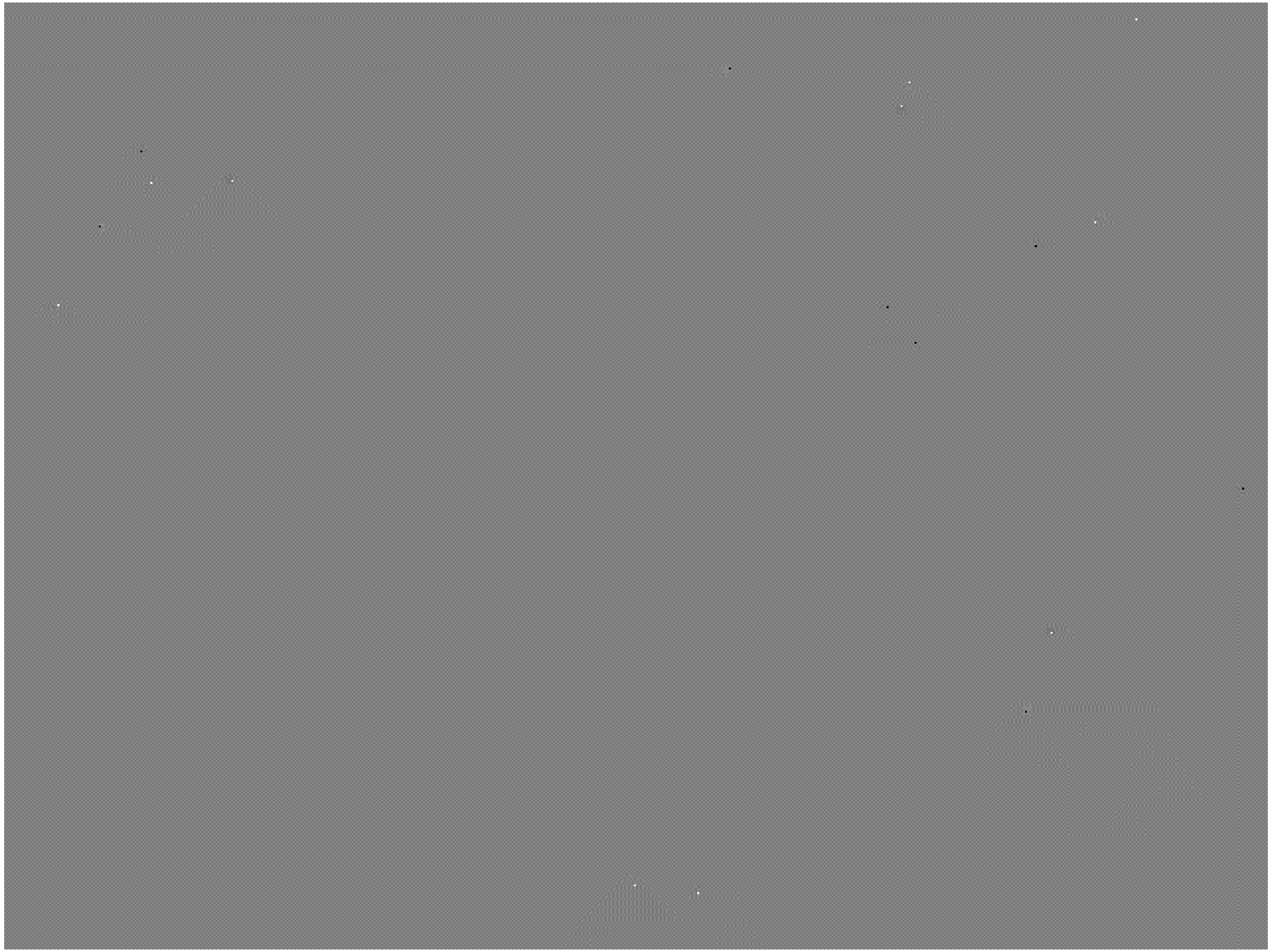
We have achieved high early classification accuracy with proposed loss.

Multimodal datasets are publicly available at https://clear-nus.github.io/visuotactile/

NUS National University of Singapore

School of Computing

Event-Based Camera

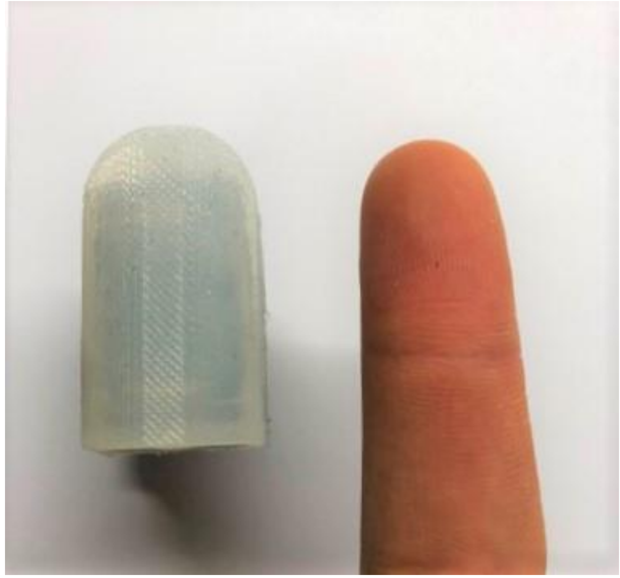


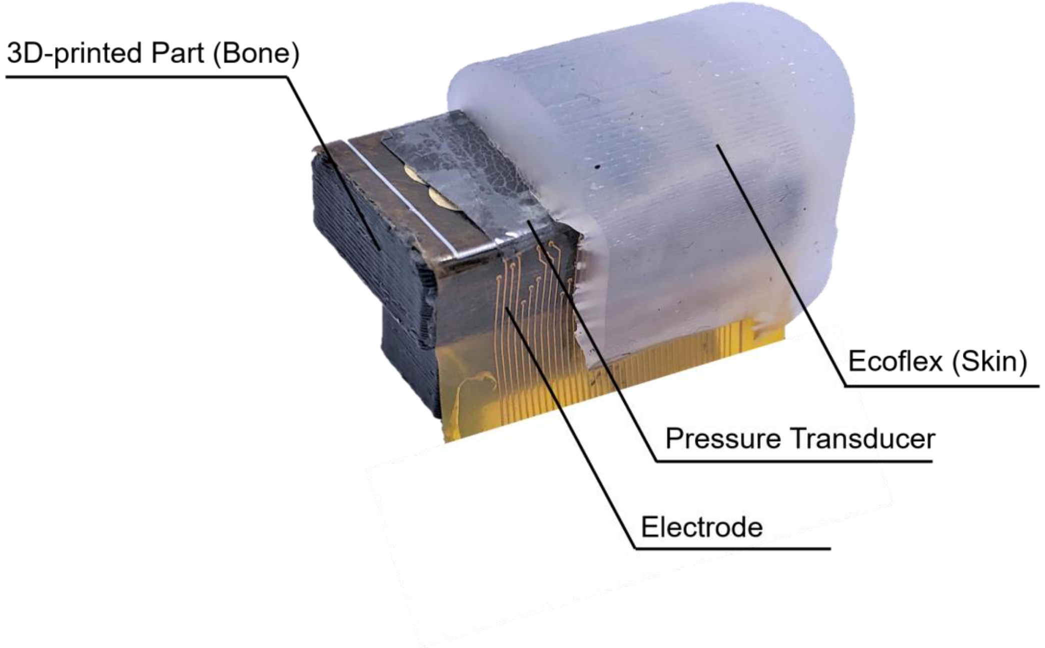
NUS National University of Singapore

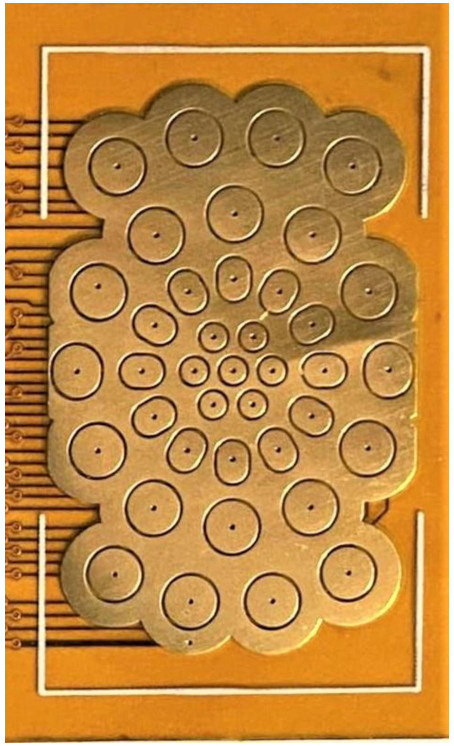
School of Computing

NeuTouch

ROBOTICS SCIENCE AND SYSTEMS



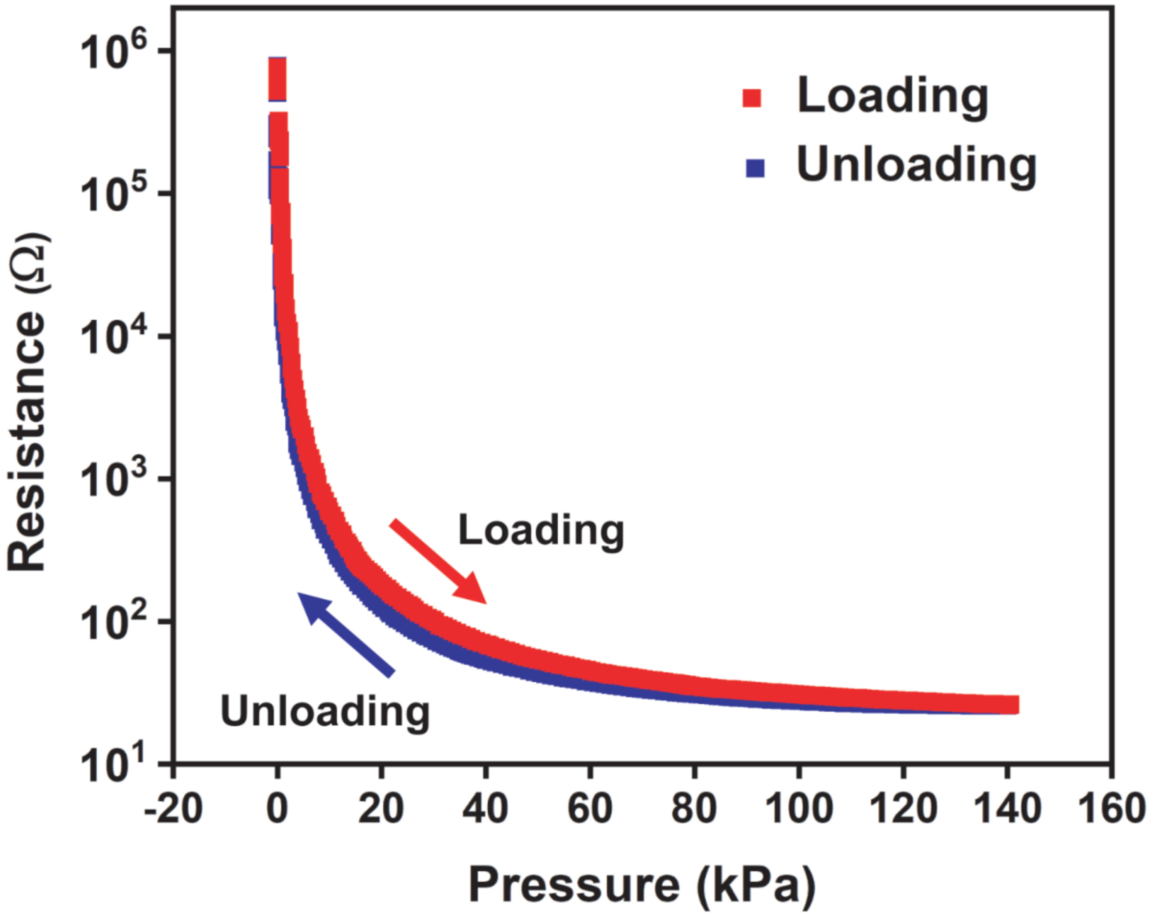




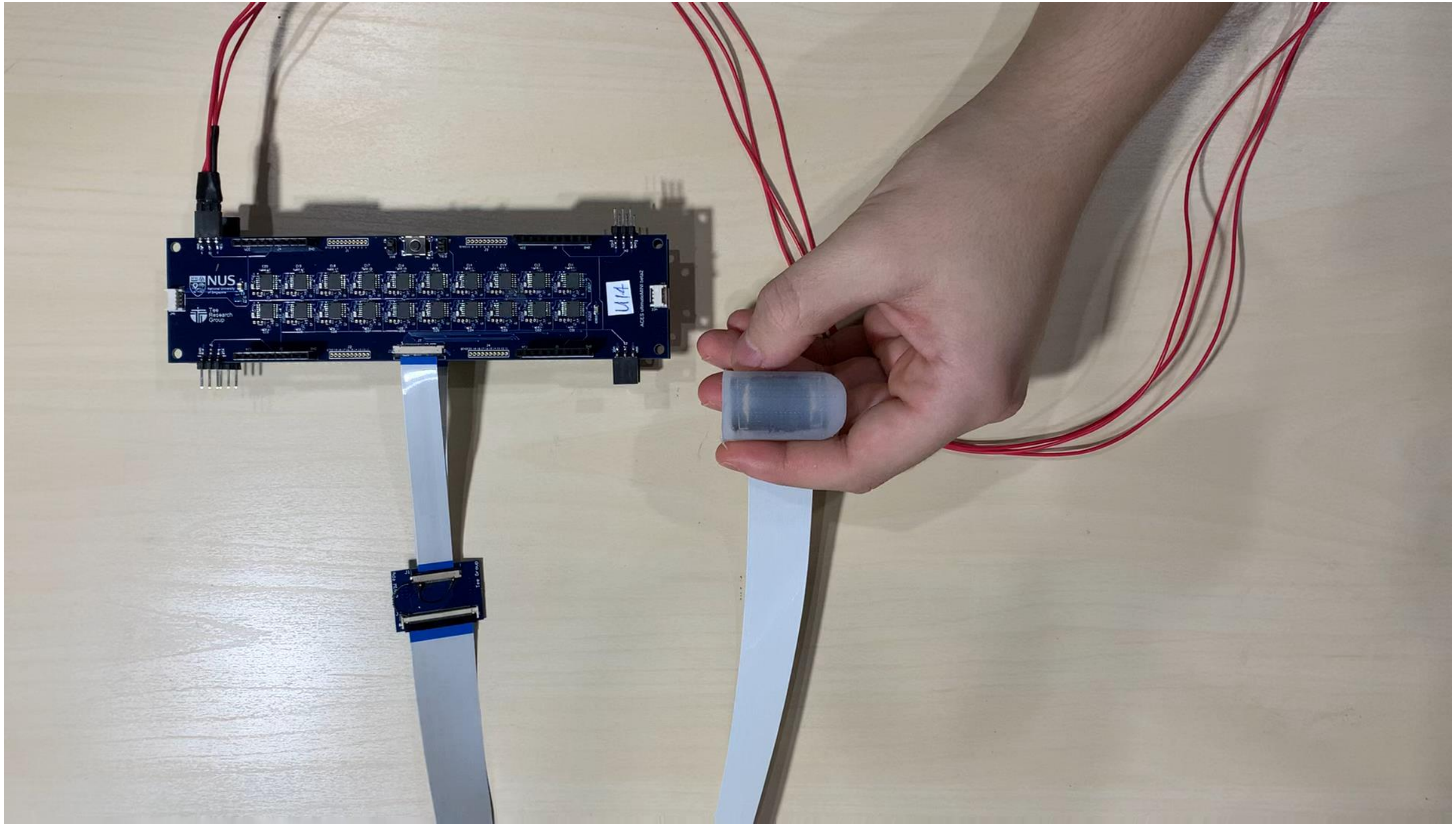
School of Computing

NeuTouch

Pressure response curve of the graphene-based pressure transducer.



NeuTouch (ACES)



Visual-Tactile Spiking Neural Network (VT-SNN)

Preprocessing:

Raw spike train

Threshold

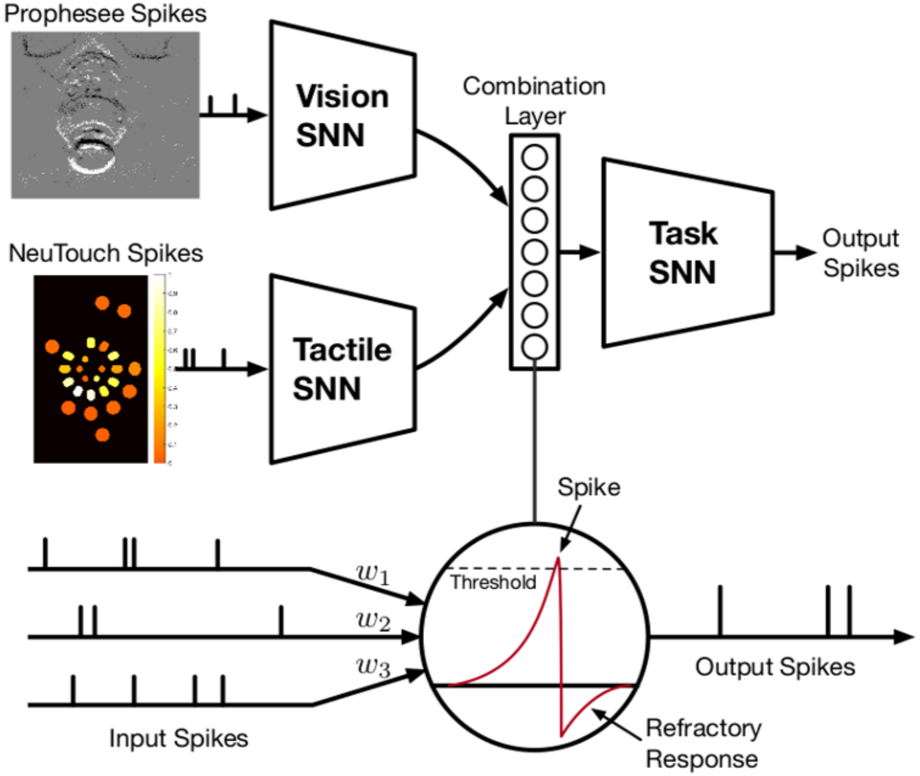
$$V_{\boxed{w}} = \begin{cases} 1 & \sum_{w} \boxed{S} \geq \boxed{S_{min}} \\ 0 & \text{otherwise} \end{cases}$$

Binning window length

Neuron Model (Spike Response Model):

$$u(t) = \sum w_i (\epsilon * s_i)(t) + (\nu * o)(t)$$

An output spike is generated whenever $u(t)$ reaches a predefined threshold ϑ.



Visual-Tactile Spiking Neural Network (VT-SNN)

Training:

Our VT-SNN is trained using SLAYER[1] framework.

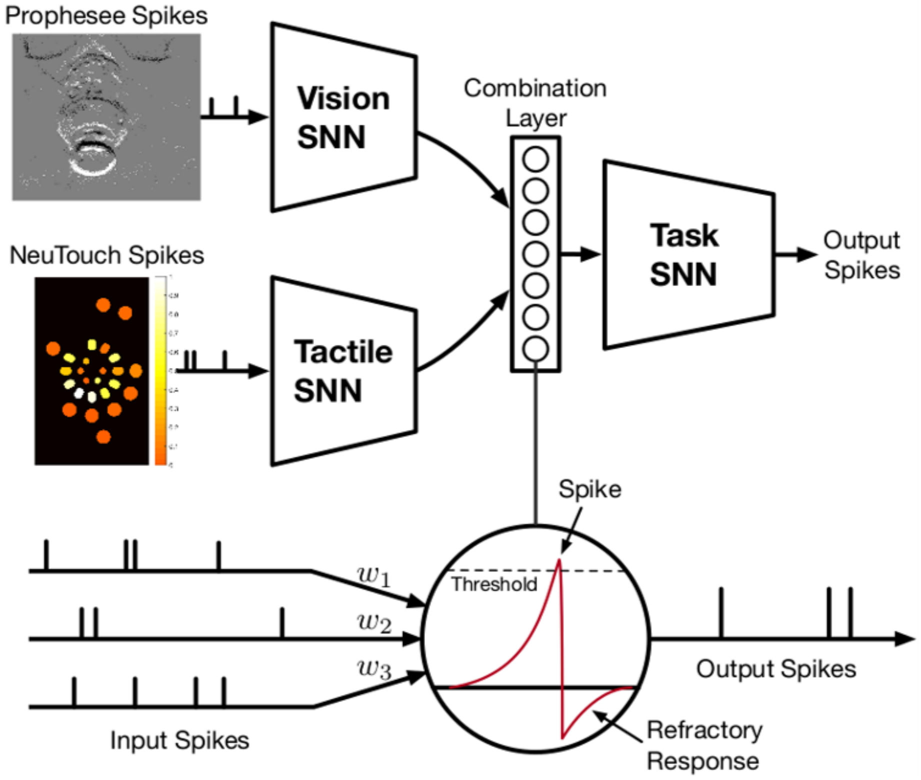
Losses

1. Standard Spike Count Loss:

$$L = \frac{1}{2}\sum_{n=0}^{N_o}\left(\underbrace{\sum_{t=0}^{T} s^n(t)}_{\text{Obtained counts}} - \underbrace{\sum_{t=0}^{T} \tilde{s}^n(t)}_{\text{Desired counts}}\right)^2$$

2. Proposed Weighted Spike Count Loss:

$$L_\omega = \frac{1}{2}\sum_{n=0}^{N_o}\left(\sum_{t=0}^{T} \omega(t) s^n(t) - \sum_{t=0}^{T} \omega(t) \tilde{s}^n(t)\right)^2$$



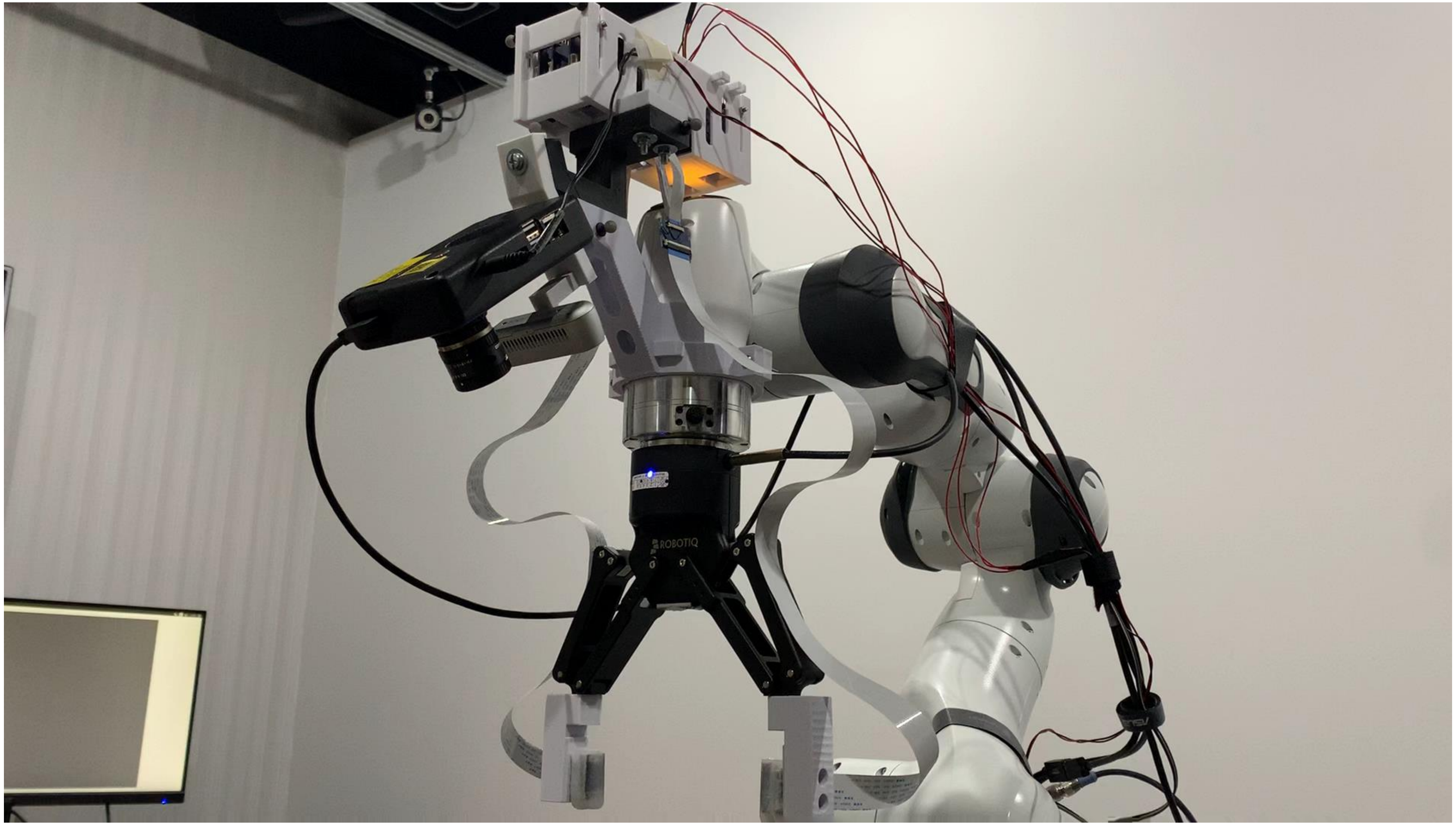
[1]. Shrestha, Sumit Bam, and Garrick Orchard. "Slayer: Spike layer error reassignment in time." *Advances in Neural Information Processing Systems*. 2018.

Visual-Tactile Spiking Neural Network (VT-SNN)

We tested VT-SNN framework on two robotic tasks:

- Container and weight classification
- Rotational slip classification

Robotic Setup

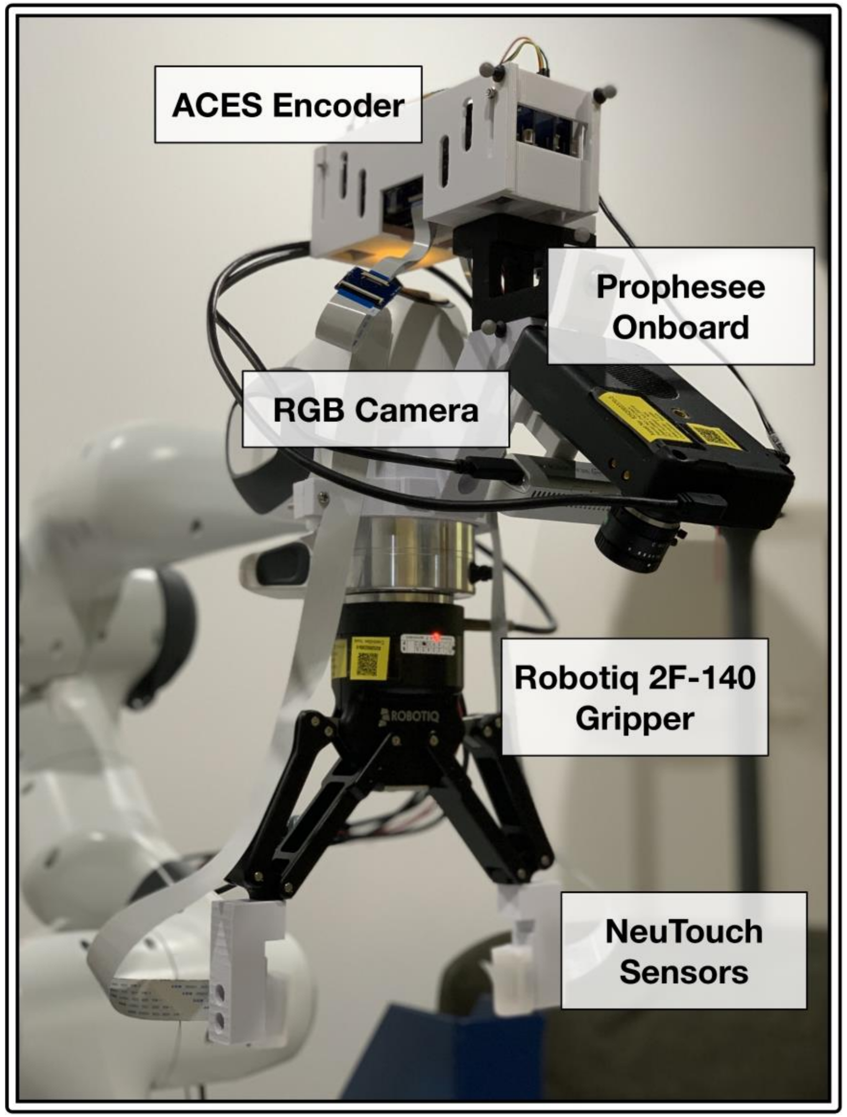


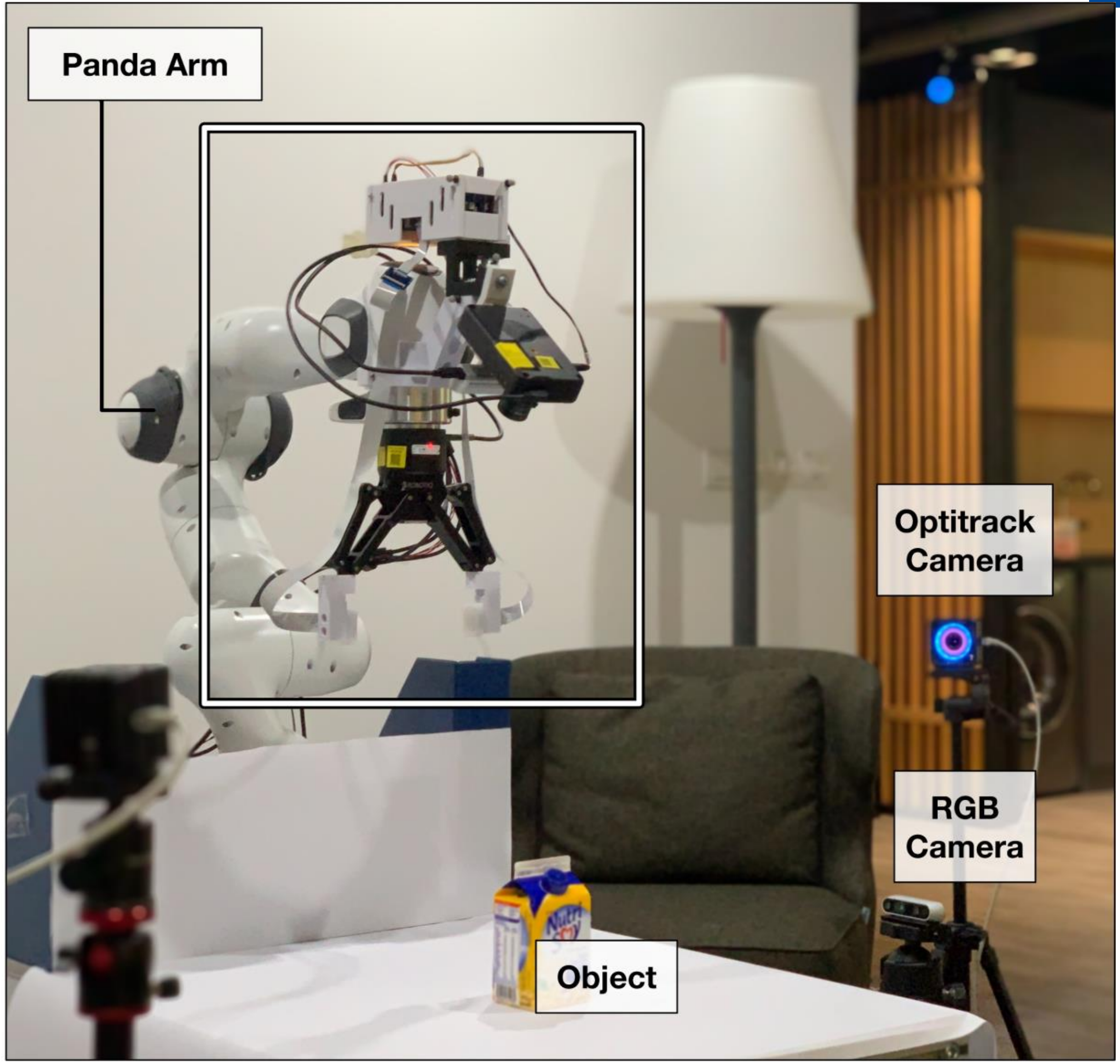
NUS National University of Singapore

School of Computing

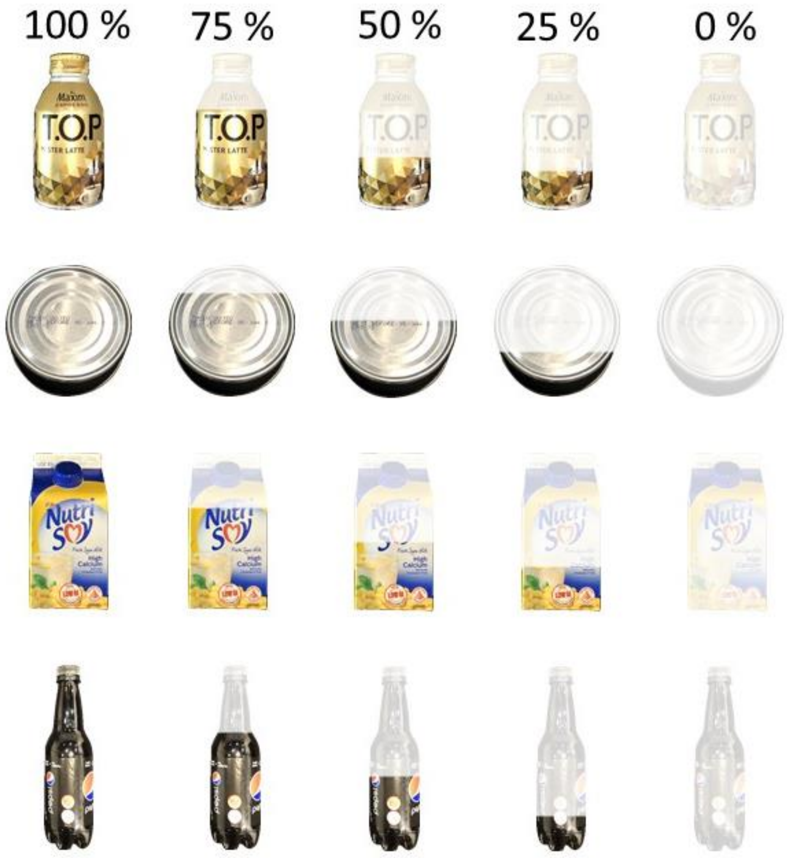
Visual-Tactile Spiking Neural Network (VT-SNN)

ROBOTICS SCIENCE AND SYSTEMS





Container and Weight Classification



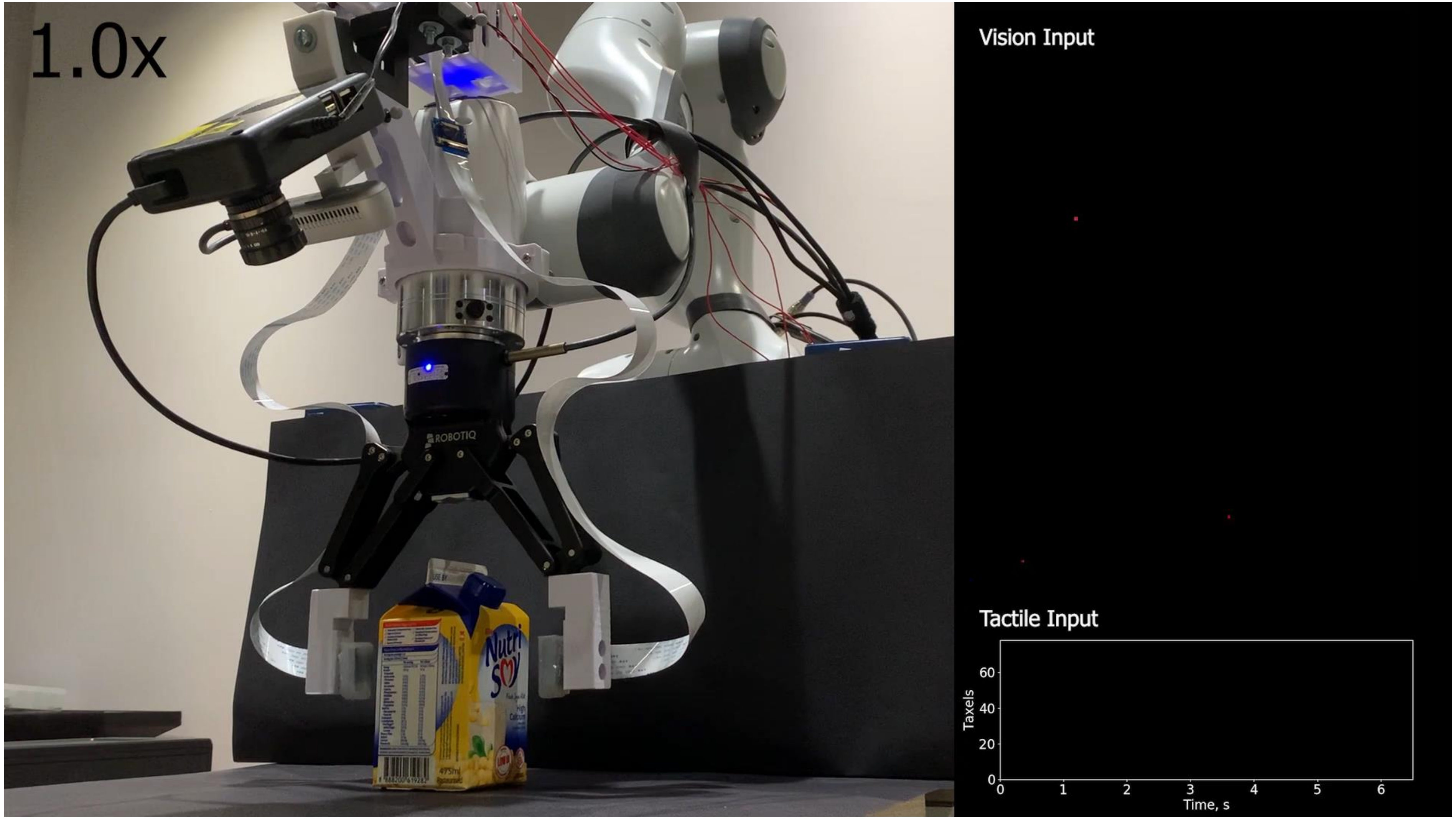
- 4 objects, 5 different weights = 20 classes
- 15 samples per each class = 300 data points
- 5-fold cross validation for testing the models

NUS
National University of Singapore

School of Computing

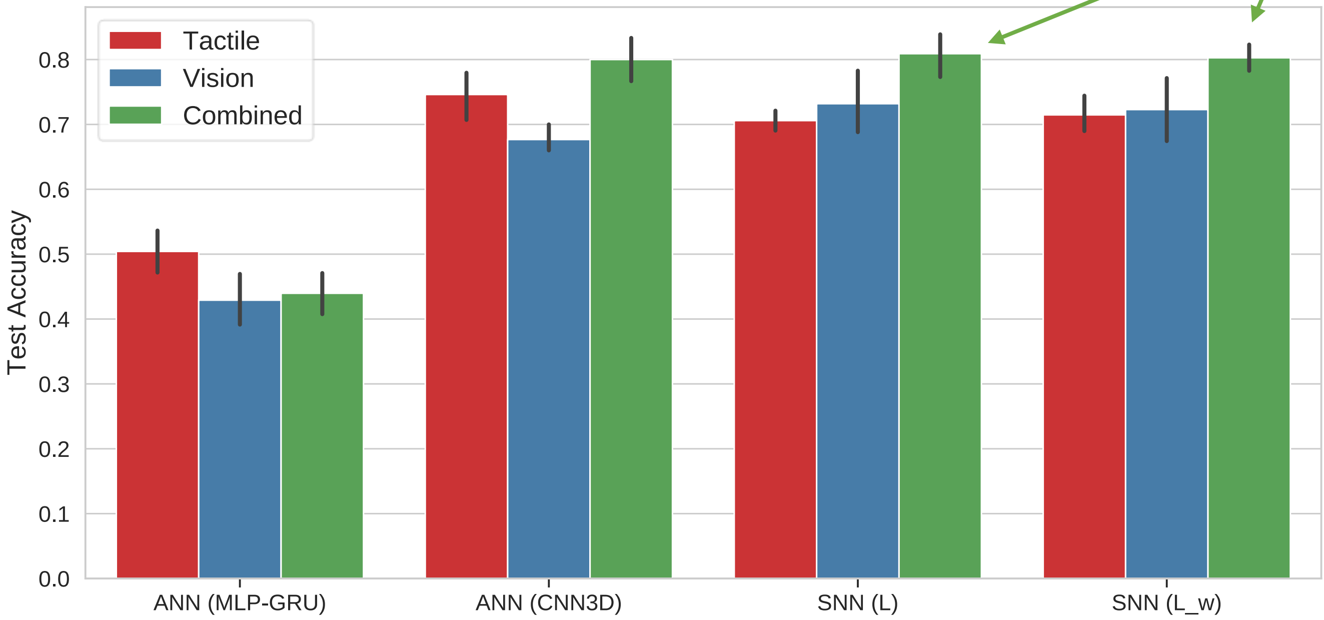
ROBOTICS
SCIENCE AND SYSTEMS

Container and Weight Classification: Data Collection Setup



Container and Weight Classification: Results (Final Accuracy)

Multimodal gives higher accuracy



SNN models perform better or comparable to ANN models

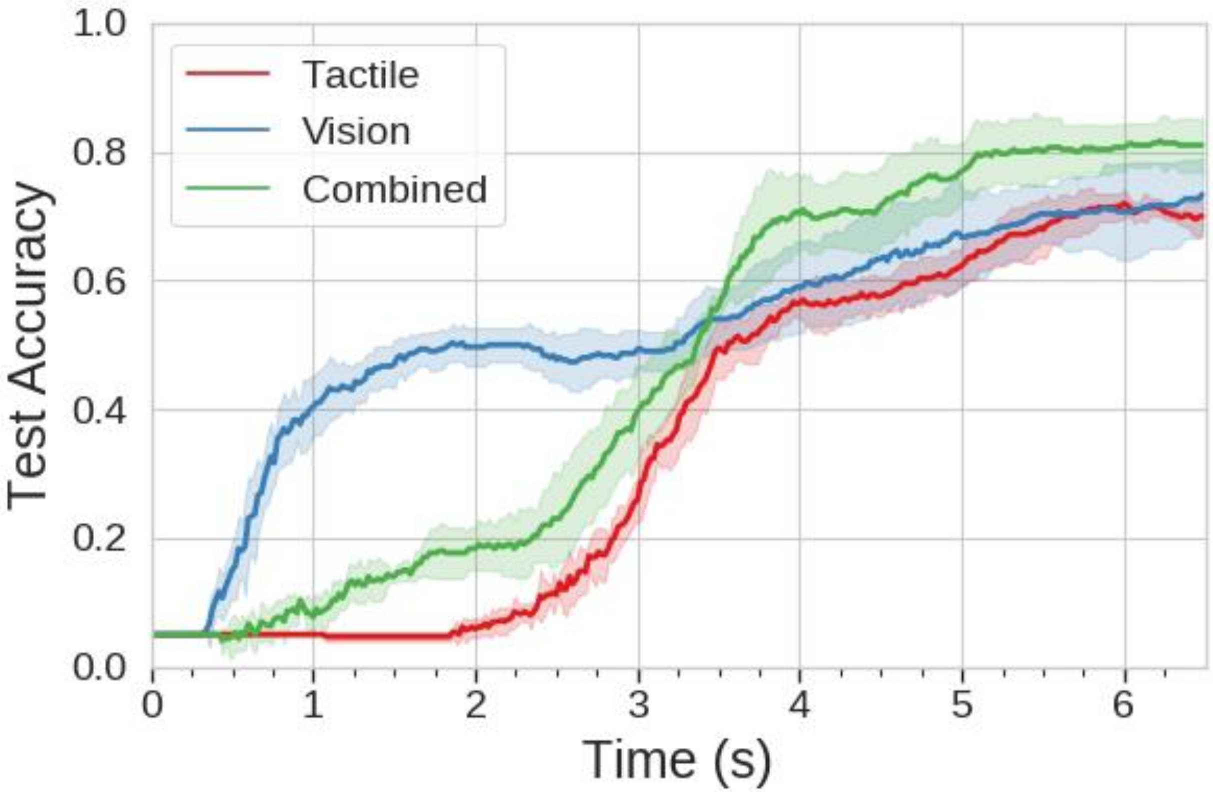
NUS National University of Singapore

School of Computing

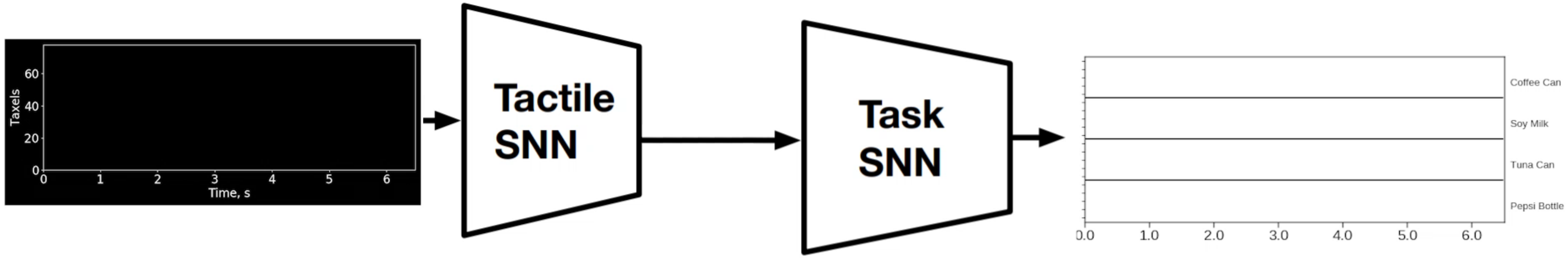
ROBOTICS SCIENCE AND SYSTEMS

Container and Weight Classification: Results (Early Classification)

$$L = \frac{1}{2}\sum_{n=0}^{N_o}\left(\sum_{t=0}^{T} s^n(t) - \sum_{t=0}^{T} \tilde{s}^n(t)\right)^2$$

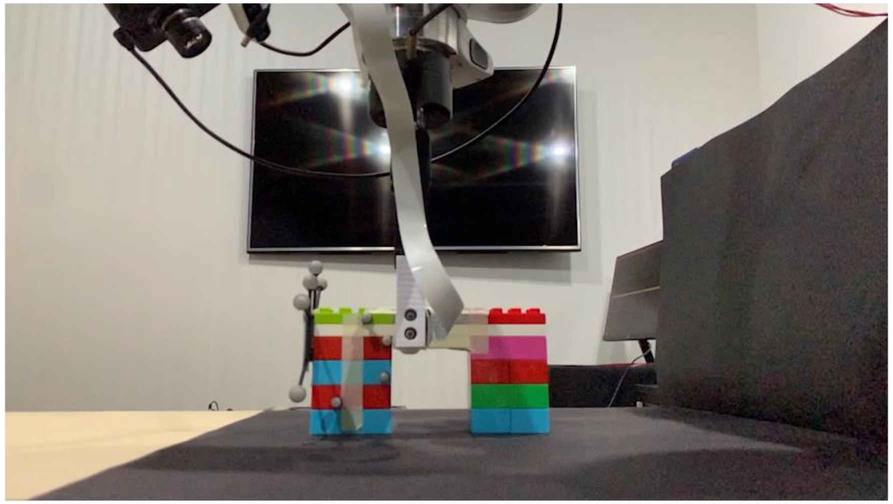


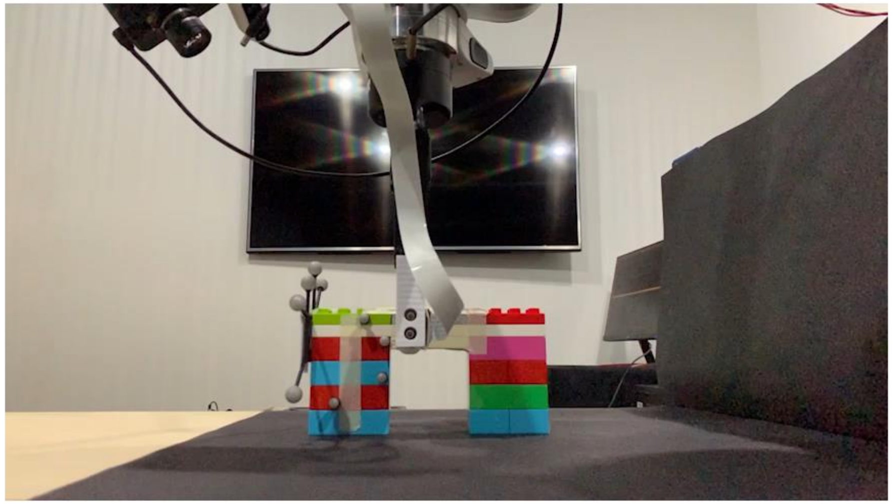
Container and Weight Classification: Output Example

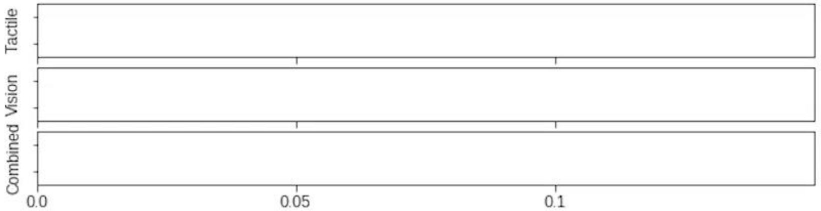


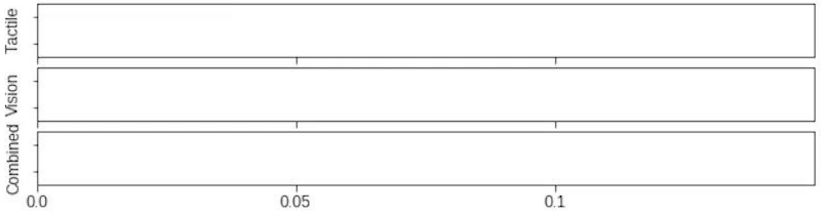
Rotational Slip Detection: Example of Output Neuron

0.02x



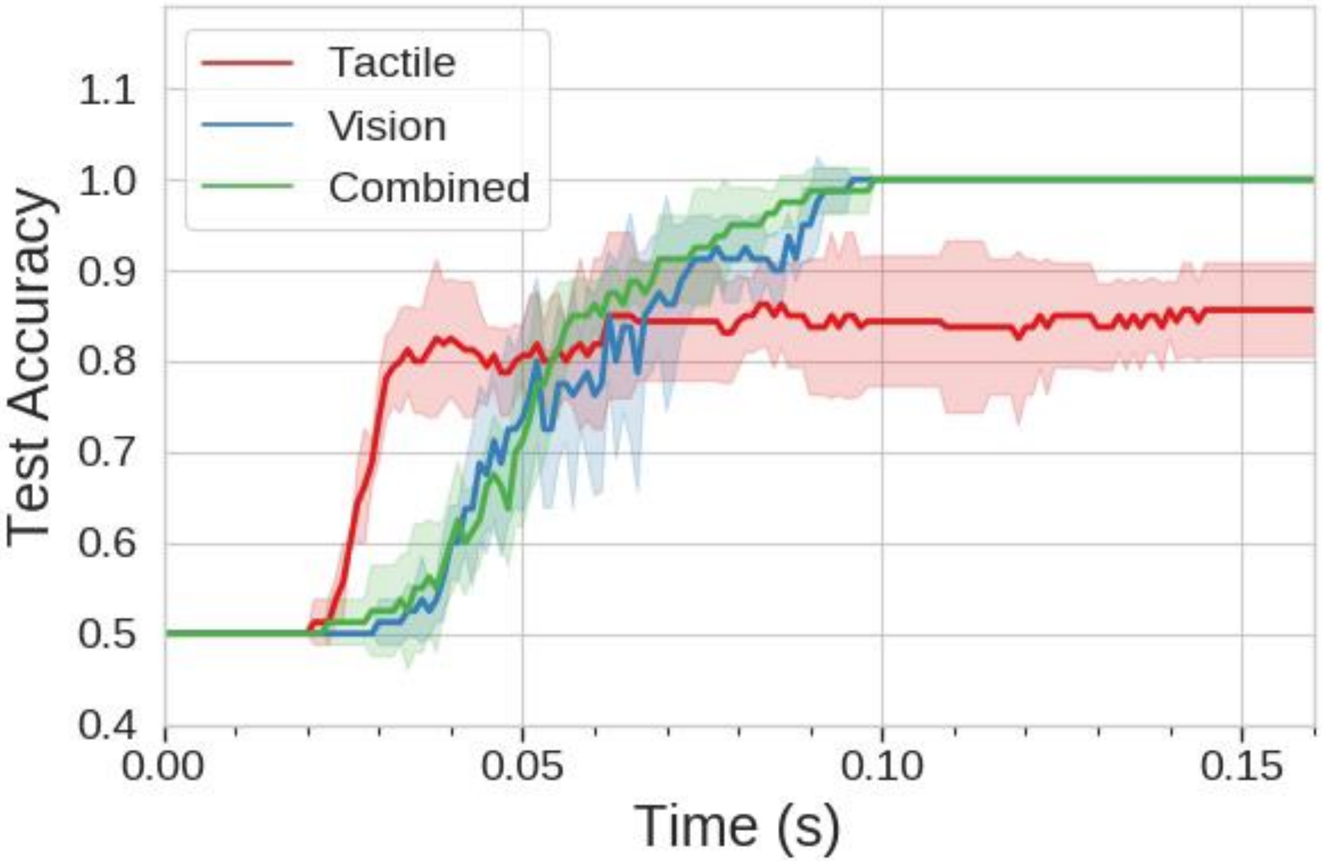






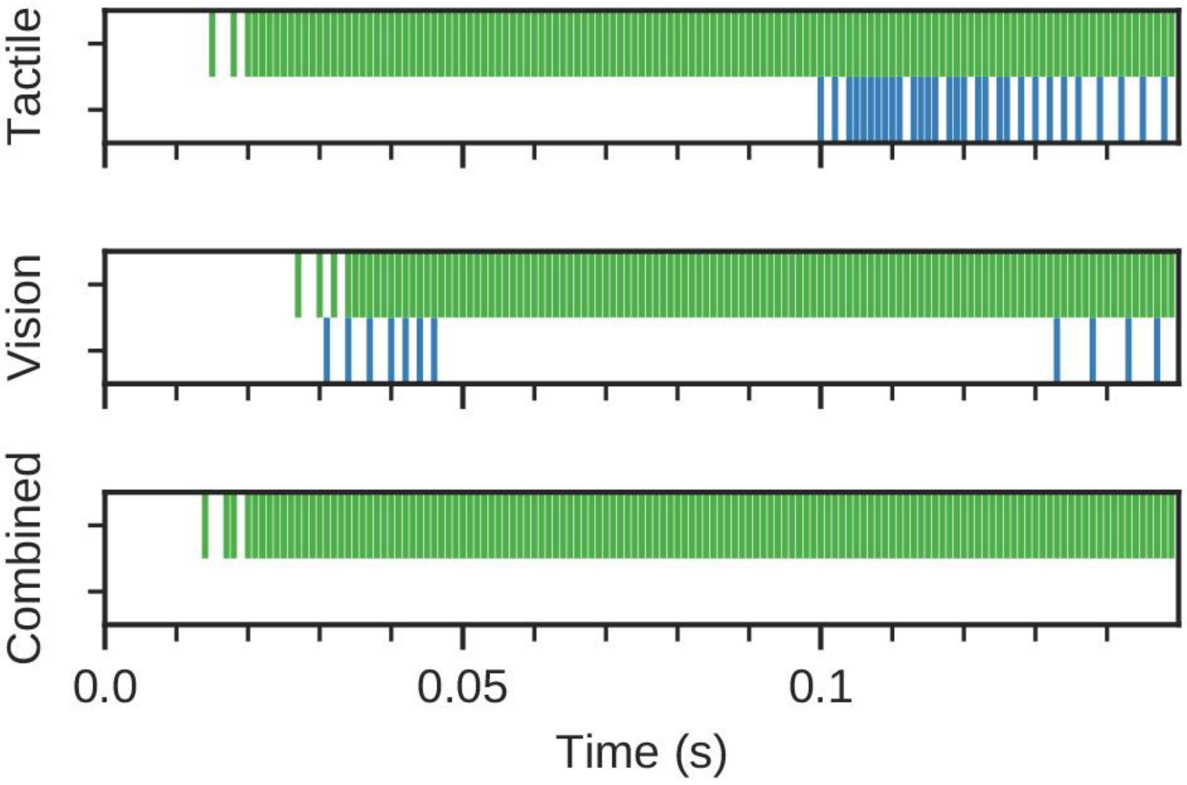
Rotational Slip Detection: Results (Early Classification)

$$L = \frac{1}{2}\sum_{n=0}^{N_o}\left(\sum_{t=0}^{T} s^n(t) - \sum_{t=0}^{T} \tilde{s}^n(t)\right)^2$$



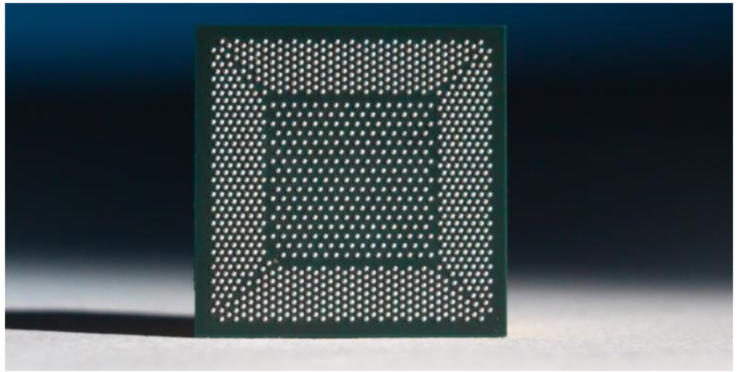
Rotational Slip Detection: Output Example with Weight Spike Count Loss

$$L_\omega = \frac{1}{2}\sum_{n=0}^{N_o}\left(\sum_{t=0}^{T}\omega(t)s^n(t) - \sum_{t=0}^{T}\omega(t)\tilde{s}^n(t)\right)^2$$

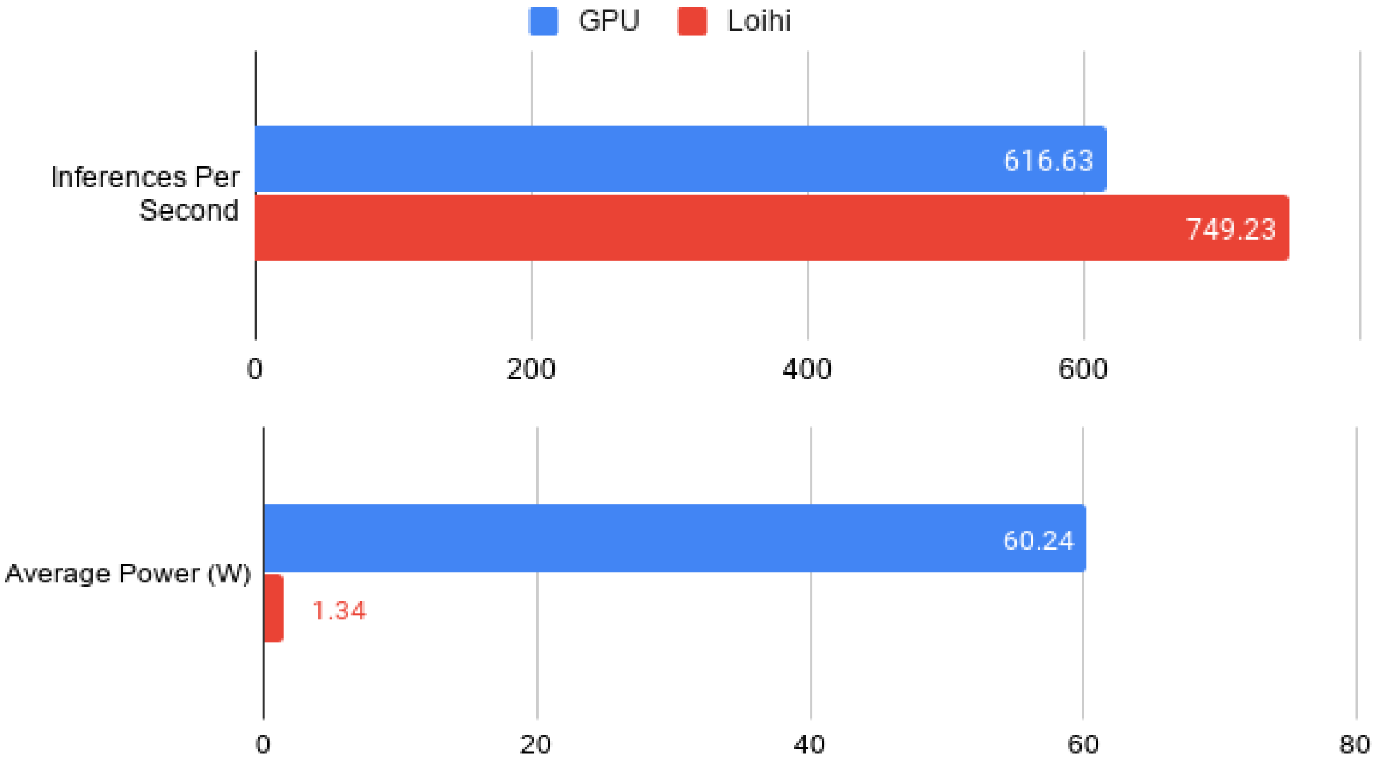


Power Utilization

Nvidia RTX 2080 Ti



Intel Loihi



NUS National University of Singapore

School of Computing

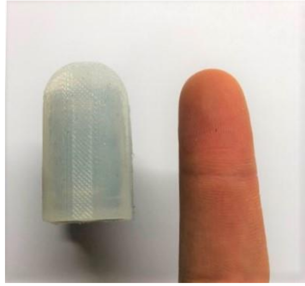
ROBOTICS SCIENCE AND SYSTEMS

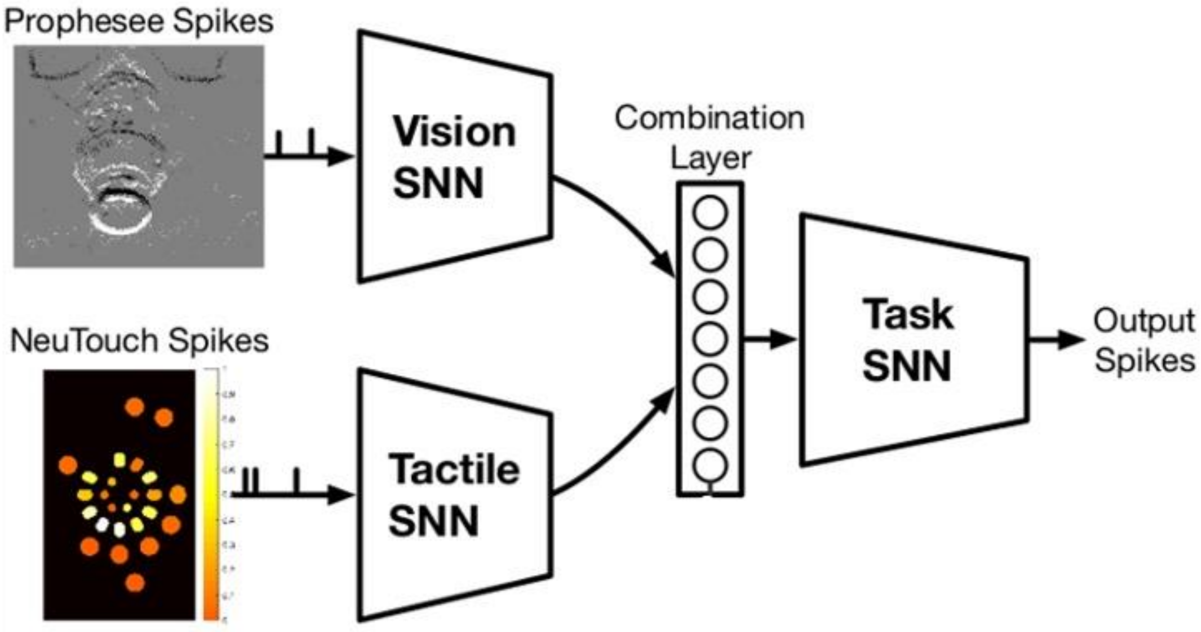
Conclusion

We proposed

- **NeuTouch** – event-based tactile sensor.
- **A VT-SNN** framework that combines side and touch event-based data.
- VT-SNN framework is tested on **two robot tasks**:
  1. Container and weight classification
  2. Rotational slip classification
- We put our datasets/code publicly available and can be found with following link:

  https://clear-nus.github.io/visuotactile/




Prophesee Spikes
Vision SNN
Combination Layer
NeuTouch Spikes
Tactile SNN
Task SNN
Output Spikes