L3S
Trustworthy AI
& Digital Transformation

Optimizing BioTac Simulation for Realistic Tactile Perception

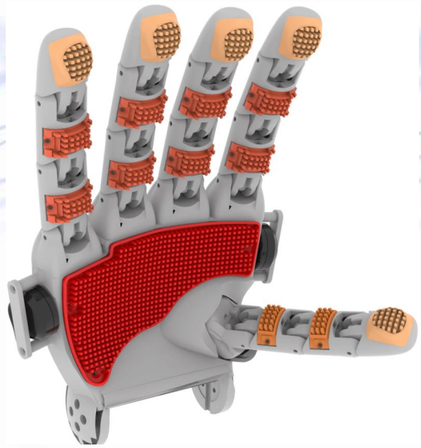
Wadhah Zai El Amri & Nicolás Navarro-Guerrero

International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, 2024.

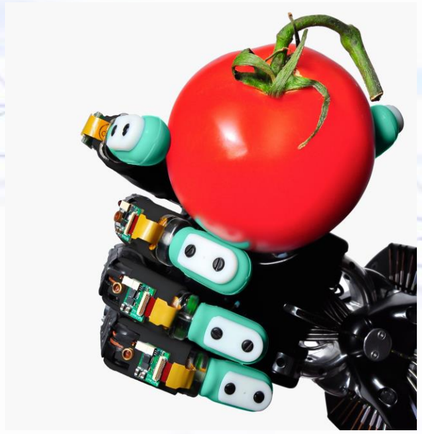
Tactile Sensing Sensors

- offer robots valuable information.
- complement knowledge coming from other modalities.

Time-series-based tactile sensors



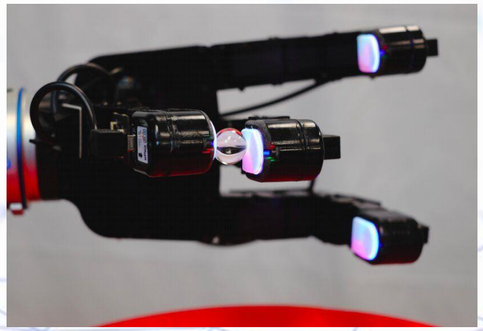
Adapted from: source



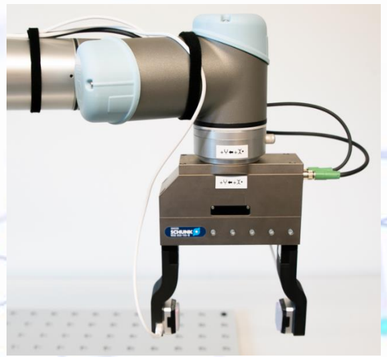
(a) (b)

Fig. 1: Time-series-based tactile sensors: (a) RH8D hand with fingerprint patterns. (b): Shadow hand with BioTac sensors.

Vision-based tactile sensors



Adapted from: source



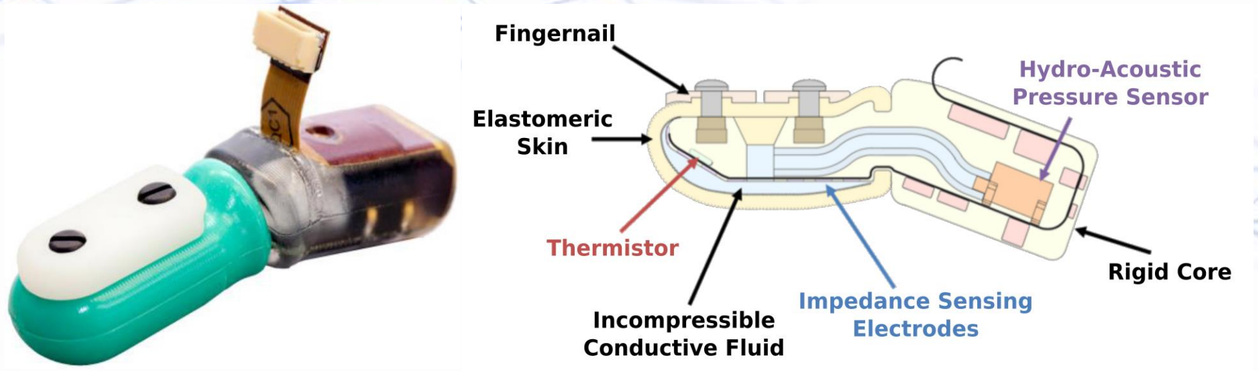
Adapted from: source

(a) (b)

Fig. 2: Vision-based tactile sensors: (a) Allegro and with DIGIT sensor (b): Schunck gripper with GelSight sensors.

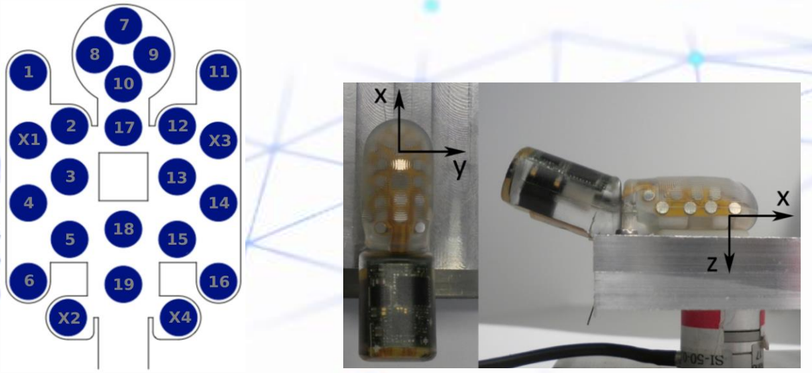
BioTac Sensor

- is a time-series-based tactile sensor.
- has a rigid core, covered with elastomeric skin filled with an incompressible conductive fluid.
- Has a sampling rate equal to 100 Hz.
- outputs: 19 electrode voltages e_1, .., e_{19}, absolute and dynamic fluid pressure p_{dc} and p_{ac}, temperature t_{dc} and heat flow t_{ac}.



(a) (b)

Fig. 3: (a): BioTac 2P Sensor [1]. (b): Internal Section of the BioTac Tactile Sensor [1].



(a) (b)

Fig. 4: (a): Map of the electrodes on the BioTac sensor [2]. (b): BioTac sensor and the position/ orientation of the coordinate frame [2].

Motivation

- Sensor is unavailable.
- Experiment repetitions are costly.

→ Simulate the BioTac sensor.

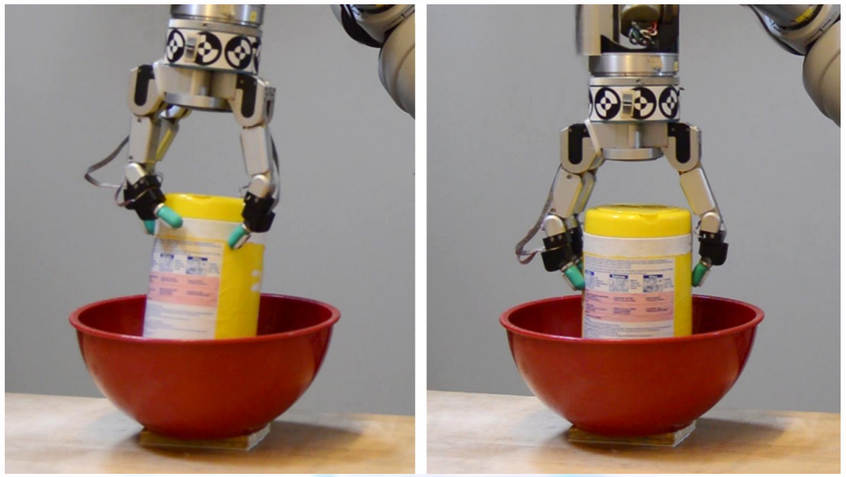


Fig. 5: Dataset collection procedure [3].

Adapted from: source

- The simulation should be:
  - Reliable.
  - Accurate
  - Real-time capable.

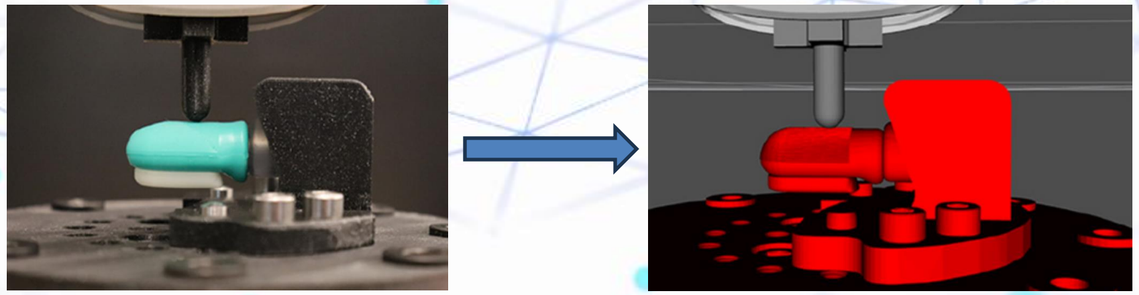
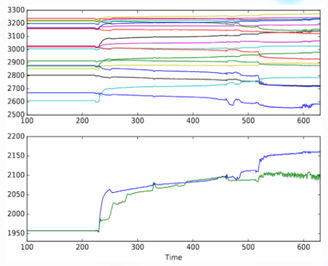


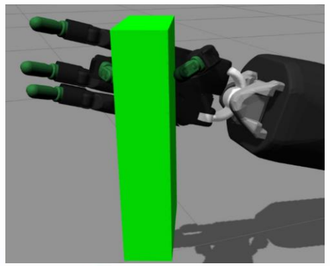
Fig. 6: Real setup and simulation [5].

Existing BioTac Sensor Simulations

Ruppel et al. [4]

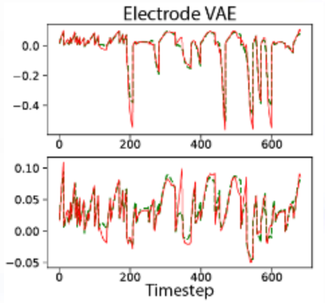
- Based on rigid-body simulation.
- A feed-forward neural network that predicts the sensor outputs.
- Requires using the temperature values in the input vector.

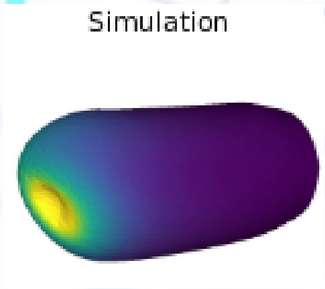




Narang et al. [5]

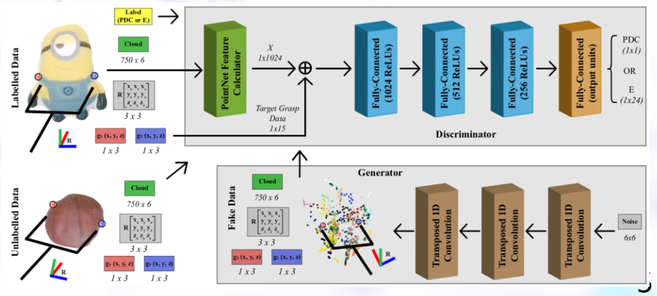
- Based on soft-body simulation.
- Uses a VAE to predict the electrode outputs from the deformation mesh.
- Pressure and vibration signals are not recorded.
- Can not be used in real-time systems.





Zaoata-Impata and Gil [6]

- Uses vision to estimate the electrode values.
- Uses a modified Semi Regression Generative Adversarial Network (SR-GAN).
- Small dataset (4000 samples).
- Vibration signals are not recorded.



Investigating Ruppel's Simulation [4]

Used Dataset:

- ca. 1h of interaction with an indenter with a spherical tip of radius equal to 2 mm, attached to a calibrated force-torque sensor.
- Approximately 290K BioTac readings with the corresponding force values and 3D indenter-tip's positions.

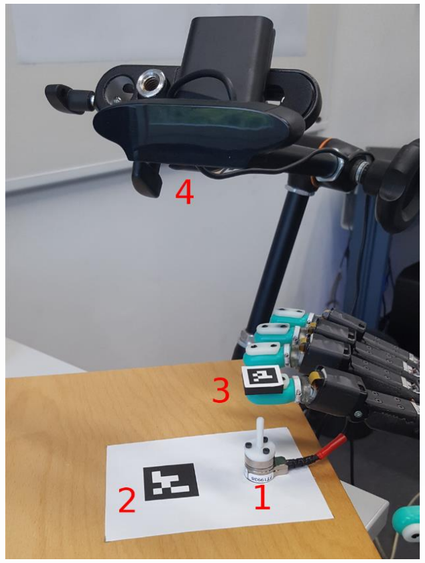


Fig. 7: Experiment setup [4].

Trained Network:

- Input Vector:
  - Position: **x**, **y**, and **z** at $t = T$,
  - Forces: $\mathbf{F_x}$, $\mathbf{F_y}$, and $\mathbf{F_z}$ at T, $T-10$, and $T+10$,
  - BioTac Temperature readings at t = T.

- Three separate dense layer columns that merge into a dense layer.

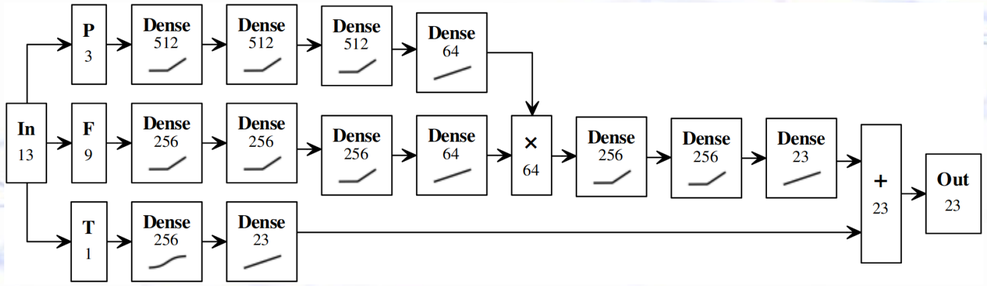


Fig. 8: Deep neural network architecture for simulating BioTac sensor outputs [4].

Investigating Ruppel et al. Simulation [4]

Investigating the Training with Temperature Readings

- Temperature values are not present in simulation:
  - Ruppel et al. suggest fixing it to the mean temperature value, when deploying.
- We conduct 10-fold cross-validation:

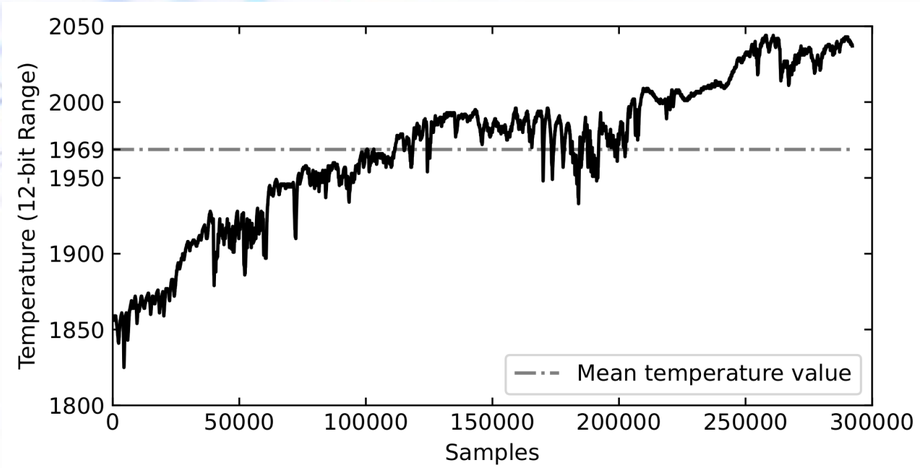


Fig. 9: Temperature values of the BioTac sensor in the Ruppel et al. [4]'s dataset.

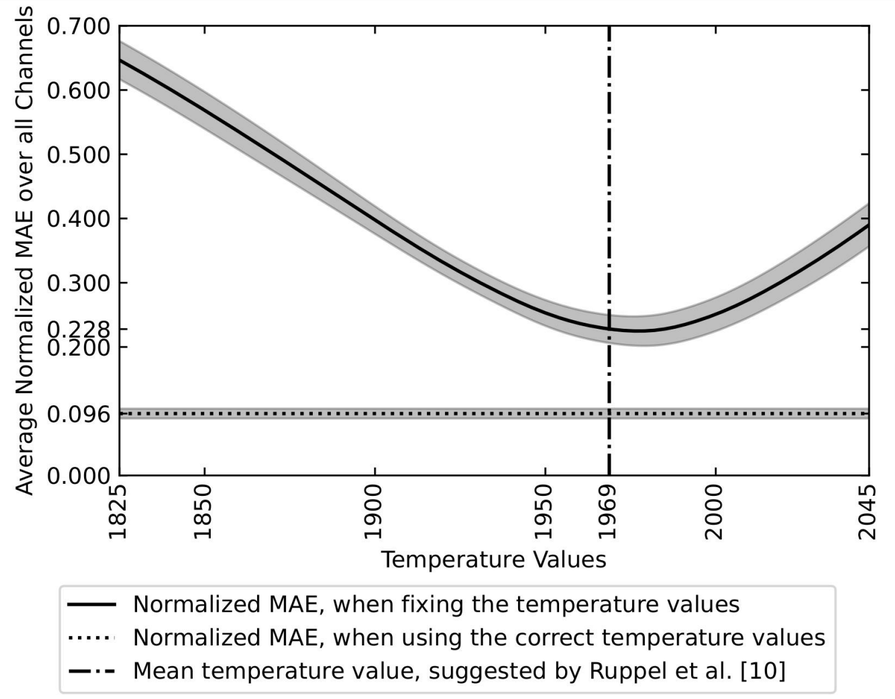


Fig. 10: Average normalized MAE over all channels for ten folds for the baseline Network.

Experiments

Training without Temperature Readings

- We propose three different networks:
- The input vector includes:
  - Position: **x**, **y**, and **z** at $t = T$,
  - Forces: $\mathbf{F_x}$, $\mathbf{F_y}$, and $\mathbf{F_z}$ at T, $T-10$, and $T+10$.
- The output vector includes:
  - 19 electrode voltages $\boldsymbol{e_1}, .., \boldsymbol{e_{19}}$ at $t = T$,
  - Absolute and dynamic fluid pressure: $\boldsymbol{p_{dc}}$ and $\boldsymbol{p_{ac}}$ at $t = T$.

XGBoost Regressor

- Fast and suitable for predicting continuous output variables.
- We use multi-regressors: one for each output channel.
- SMAC3 [8] is used to find the best hyperparameters.

Feed-Forward Neural Network

- Commonly used in regression tasks.
- SMAC3 [8] is used to find the best hyperparameters.

Transformer Encoder

- Proved beneficial in several time-series tasks.
- Based on ViT encoder.
- SMAC3 [8] is used to find the best hyperparameters.

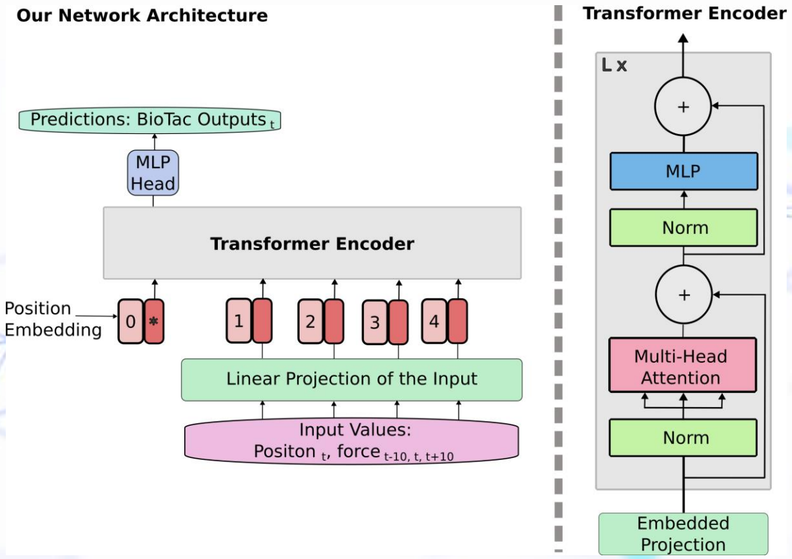


Fig. 11: Visualization of our used transformer architecture, based on Dosovitskiy et al. [9].

Experiments

Training without Temperature Readings

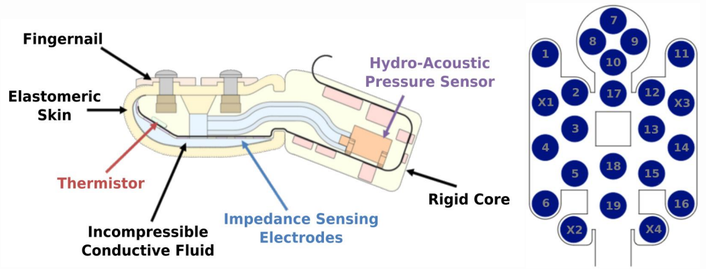
- XGBoost Regressor achieves the lowest error.
- The transformer encoder achieves comparable results with a low number of parameters.

Table 1: Results of all trained network over all ten folds.

	Nb. Param.	MAE	Norm. MAE	MAE	Norm. MAE
		Over all Channels		Electrodes Only	
Ruppel et al. [10]	806K	18.977 (1.719)	0.228 (0.022)	17.147 (1.642)	0.232 (0.023)
Our XGBoost Regressor	1584K	**13.368** (1.340)	**0.150** (0.014)	**11.446** (1.204)	**0.150** (0.015)
Our Feed-Forward Neural Network	2233K	14.693 (1.386)	0.168 (0.015)	12.724 (1.252)	0.169 (0.016)
Our Transformer Encoder	**599K**	13.760 (1.400)	0.156 (0.016)	11.736 (1.280)	0.155 (0.016)

Experiments

Training without Temperature Readings



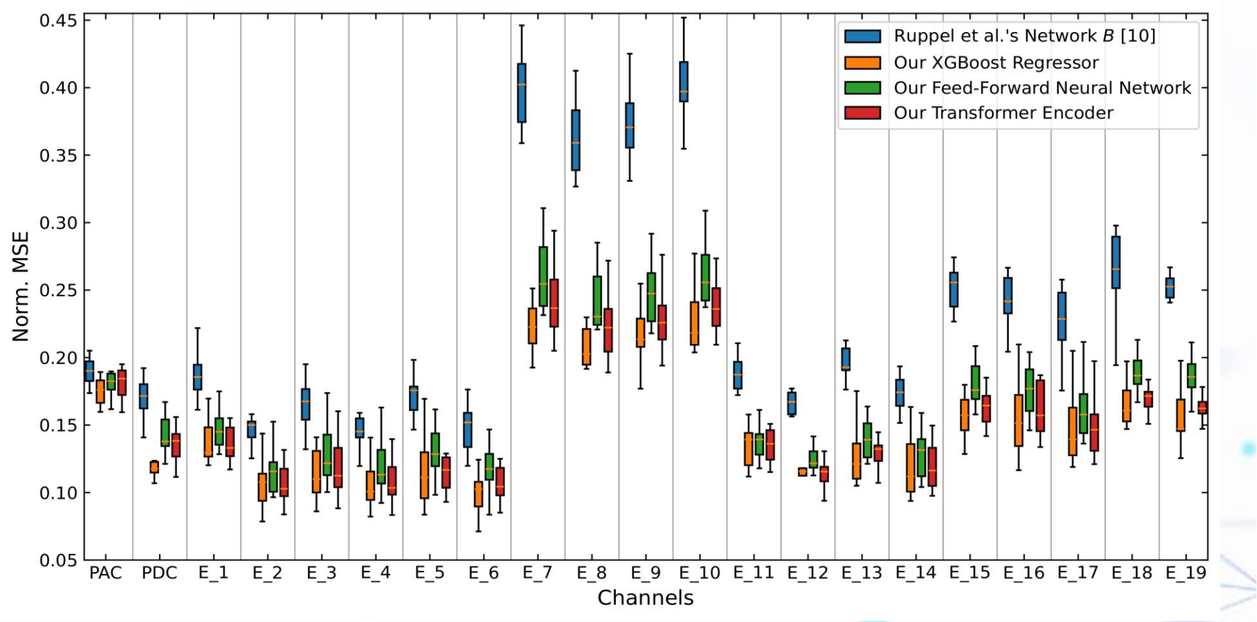


Fig. 12: Distribution of the normalized MAE for each channel and all models.

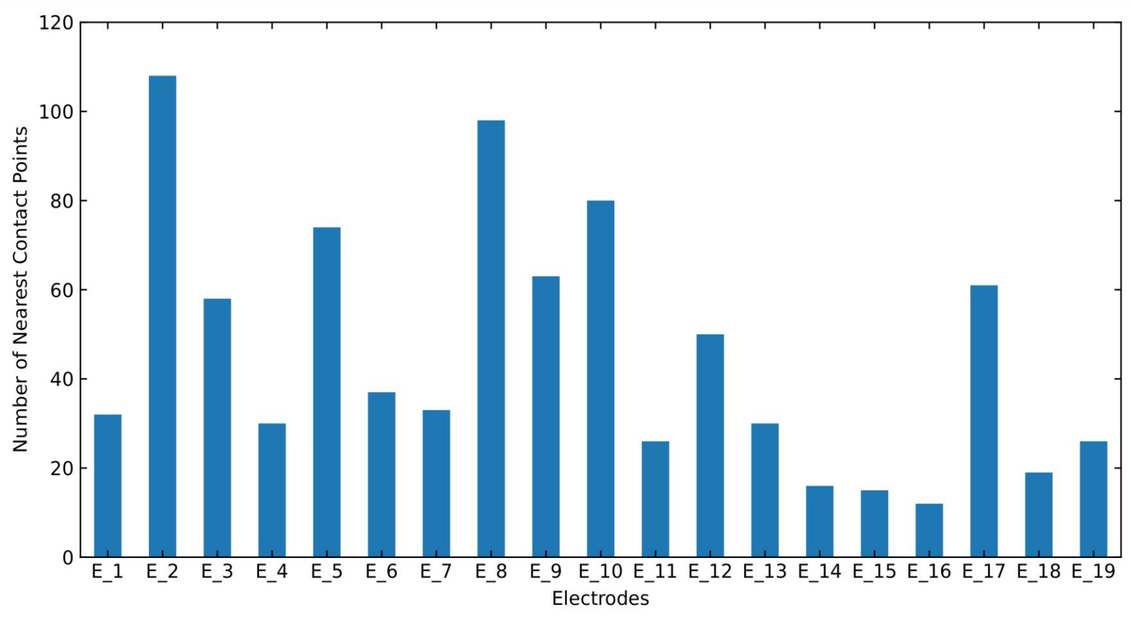


Fig. 13: Distribution of nearest contact points for each electrode.

→ Non-uniform distribution of the number of touches near each electrode, implying an unbalanced dataset.
→ The BioTac sensor does not have radial symmetry.
→ The fluid volume is not the same throughout the sensor. Particularly near electrodes 7, 8, 9, and 10.

Experiments

Varying the Windowing Size

- Investigate the following input combinations:
  1. Forces Ft: $\forall\ t \in \{T, T - 10, T + 10\}$.
  2. Forces Ft: $\forall\ t \in \{T, T - 10\}$.
  3. Forces Ft: $\forall\ t \in \{T\}$.
  4. Forces Ft: $\forall\ t \in \{T, T - 10, T - 5, T + 5, T + 10\}$.
  5. Forces Ft: $\forall\ t \in [T - 10, T + 10], t \in \mathbb{N}$.
  6. Forces Ft: $\forall\ t \in [T - 10, T], t \in \mathbb{N}$.
  7. Current, last, and next position and forces Ft: $\forall\ t \in \{T, T - 10, T + 10\}$.
  8. Current, last, and next position and forces Ft: $\forall\ t \in [T - 10, T + 10], t \in \mathbb{N}$.

Experiments

Varying the Windowing Size

- Investigate the following input combinations:
  1. Forces Ft: $\forall\, t \in \{T, T-10, T+10\}$.
  5. Forces Ft: $\forall\, t \in [T-10, T+10], t \in \mathbb{N}$.
  7. Current, last, and next position and forces Ft: $\forall\, t \in \{T, T-10, T+10\}$.
  8. Current, last, and next position and forces Ft: $\forall\, t \in [T-10, T+10], t \in \mathbb{N}$.

Table 2: Significance test with the corrected paired t-Test [7].

	5 vs 1	7 vs 1	8 vs 1
Our XGBoost Regressor	0.282% (0.999)	-0.055% (0.190)	-0.208% **(0.009)**
Our Feed-Forward Neural Network	0.217% (0.715)	0.862% (0.961)	0.402% (0.758)
Our Transformer Encoder	0.077% (0.631)	0.188% (0.672)	-0.662% **(0.047)**

→ The error improvement is small. However, it is statistically significant.

→ Including last and next position adds valuable information, when combined with shorter force intervals.

Experiments

Varying the Windowing Size

Table 3: Significance test with the corrected paired t-Test [7].

vs	Our XGBoost Our FFNN	Our XGBoost Our Transformer	Our Transformer Our FFNN
1	-1.778% **(0.000)**	-0.531% (0.074)	-1.247% **(0.002)**
2	-1.360% **(0.001)**	-0.259% (0.250)	-1.101% **(0.000)**
3	-1.272% **(0.003)**	0.851% (0.996)	-2.123% **(0.000)**
4	-1.989% **(0.000)**	-0.032% (0.459)	-1.956% **(0.000)**
5	-1.279% **(0.002)**	-0.173% (0.322)	-1.107% **(0.002)**
6	-1.106% **(0.003)**	0.092% (0.616)	-1.197% **(0.002)**
7	-2.696% **(0.000)**	-0.774% (0.105)	-1.921% **(0.000)**
8	-2.389% **(0.001)**	-0.077% (0.439)	-2.312% **(0.002)**

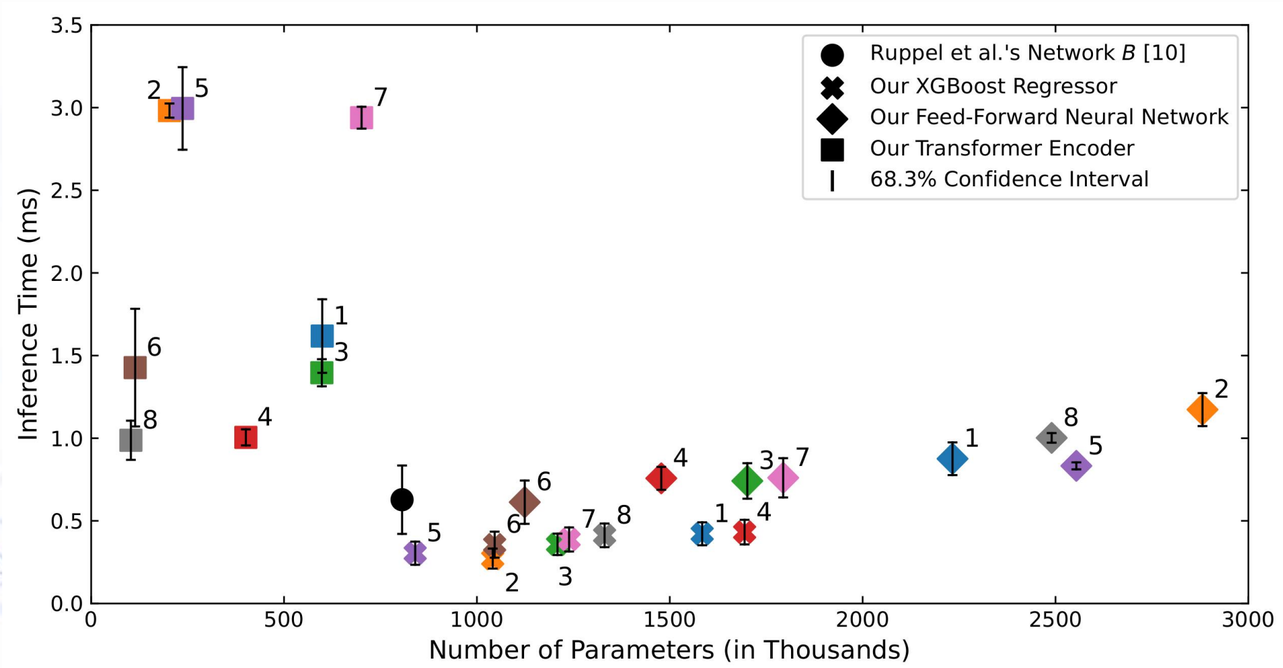


Fig. 14: Visualization of inference time against the number of parameters for each model over all input combinations.

→ There is not statistical significancy difference between XGBoost Regressor and Transformer Network.

→ Transformers have less parameters. However, they have higher inference time.

Conclusion

Our Contribution:

- We include a thorough analysis of Ruppel et al.'s work [4].
- We introduced three alternative models, that are independent of the temperature readings, resulting in an absolute error reduction of 8.0%, and a relative error improvement of 14.9%.
- We report on the limitations of the dataset. Particularly, the dataset is unbalanced and only includes a single indenter type.

Future Work:

- Extend the dataset and transfer it to a soft-body simulation, i.e., Isaac Gym.
- Transfer the BioTac dataset to other sensor modalities.
  For example: **time-series-based modality ↔ vision-based modality**.

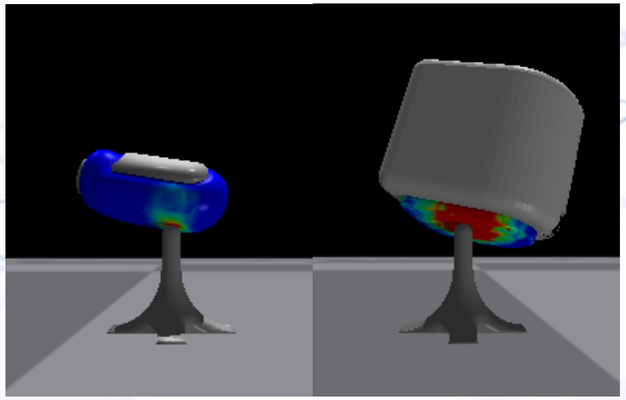


Fig. 15: Dataset collection environment for BioTac and DIGIT sensor.

Contact

Wadhah Zai El Amri

wadhah.zai@l3s.de

www.l3s.de