Geometrical Reconstruction using Acoustic Tactile Sensing

Georg Wolnik

TU Berlin - Robotics and Biology Department

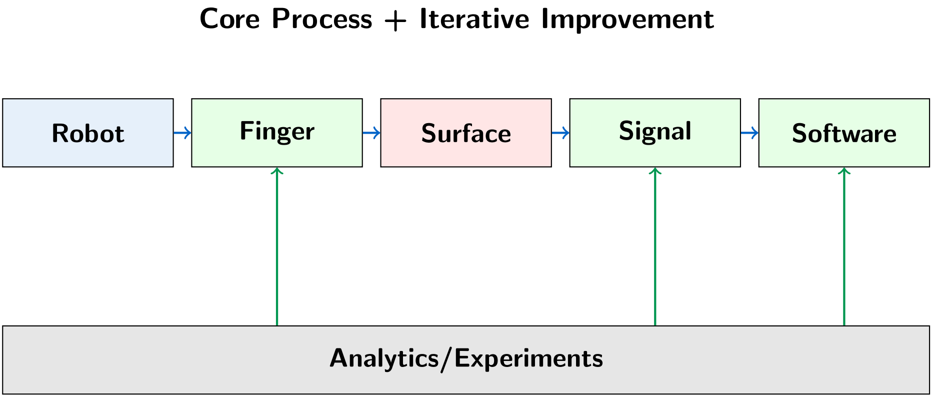
December 14, 2025

End-to-end Project Pipeline

Core Process

Robot → **Finger** → **Surface** → **Signal** → **Software**

End-to-end Project Pipeline



Research Question

How can we achieve geometrical reconstruction of complex 3D shapes using acoustic tactile sensing?

Sub-Questions:

- **Can we distinguish between different contact scenarios?**
- **Which models perform the best on our datasets?**
- **What information can be extracted from acoustic responses?**
- **What information should we focus on?**
- **Which Frequencies Matter the Most?**

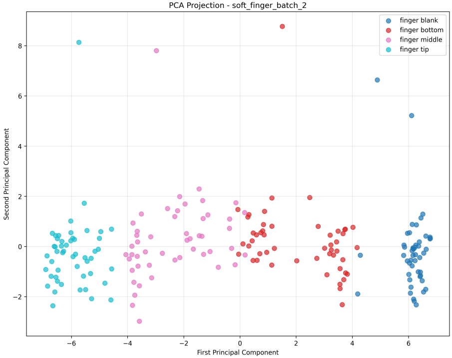
Datasets

Dataset	Contact Type	Classes	Samples
Batch 1	Position detection	4	200
Batch 2	Position detection	4	200
Batch 3	Edge detection (simple)	3	150
Batch 4	Paper clip detection	2	100
Edge v1	Edge detection (complex)	3	630
Total		**2-4**	**1,280**

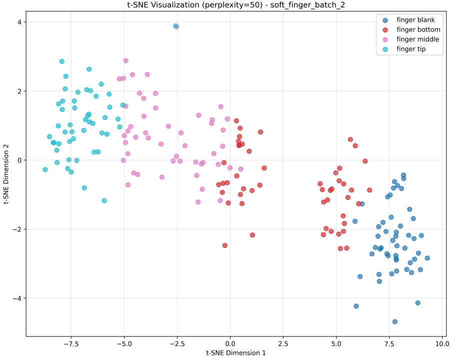
Batches 1-4 have 50 samples per class. Edge v1 has 210 samples per class.

Can we distinguish between different contact scenarios?

PCA Analysis



t-SNE Analysis



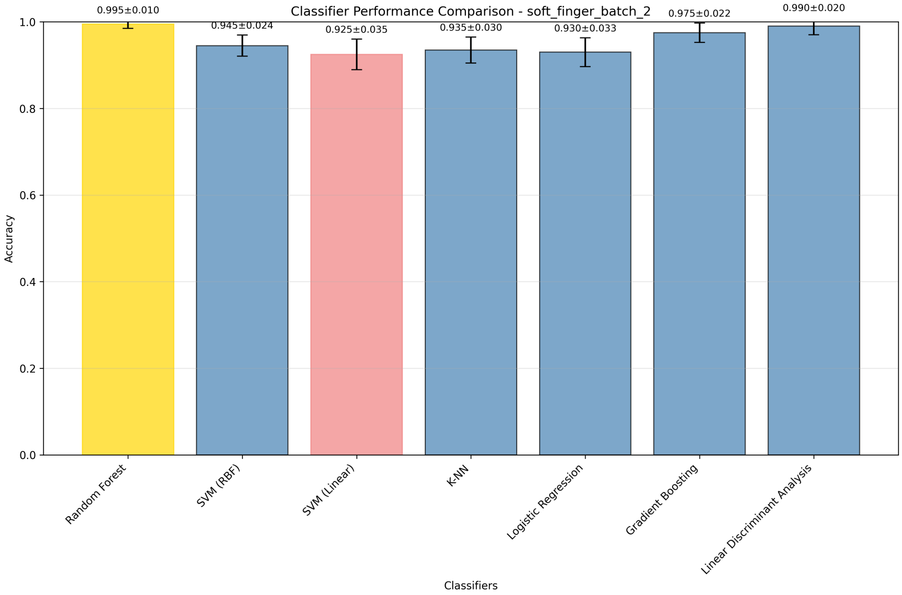
Classification Performance

Which models perform the best on our datasets?

Tested Classifiers:

- **Random Forest** – Combines many decision trees to make smarter predictions
- **SVM (RBF)** – Draws the best curved boundary to separate different classes
- **SVM (Linear)** – Draws a straight line to separate different classes
- **K-Nearest Neighbors** – Classifies based on the majority vote of nearby similar examples
- **Logistic Regression** – Predicts probabilities using a straight-line relationship
- **Gradient Boosting** – Builds trees one by one, each fixing the mistakes of the previous
- **Linear Discriminant Analysis** – Finds the best direction to tell classes apart

Model Performance - Batch 2



What information can be extracted from acoustic responses?

1. **Direct Signal Analysis**
   - **Waveform**: overall energy, fade over time
   - **Frequency Spectrum**: center, bandwidth
   - **Time-Frequency View**: freq changes, spectrum contrasts
   - **High Frequencies**: energy ratio above 8 kHz
2. **System Response Analysis (Transfer Function)**
   - **Signal Separation** $H(f) = Y(f)/X(f) \rightarrow$ true system behavior
   - **Peak Frequencies**: peak amplitude, peak frequency location
   - **Frequency Stats**: Q-factor, asymmetry

What information should we focus on?

How Feature Ablation Works:

- Remove one feature at a time and retrain
- Measure accuracy drop

Main Finding:

- Subsets of features can be used across all datasets

Dataset	Baseline	Optimal	Optimal Acc.	Max Drop	Increase Ex.
Batch 1	96%	2/38	99%	1.0%	Feature 35 (+0.5%)
Batch 2	99.5%	10/38	99.5%	1.5%	None
Batch 3	99.3%	1/38	100%	1.3%	None
Batch 4	85%	10/38	86%	2.1%	None
Edge v1	66%	13/38	71%	5.6%	Feature 11 (+0.8%)

What information should we focus on?

Dataset	Optimal Features
Batch 1	ultra_high_energy_ratio, spectral_bandwidth, ultra_high_ratio, high_energy_ratio, spectral_centroid, mid_energy_ratio, spectral_flatness, env_skew, low_mid_ratio, env_kurtosis, high_freq_decay_rate
Batch 2	spectral_centroid, ultra_high_energy_ratio, high_energy_ratio, ultra_high_ratio, low_mid_ratio, mid_energy_ratio, burst_rms, env_kurtosis, env_skew, temporal_centroid, damping_ratio
Batch 3	spectral_bandwidth, high_energy_ratio, spectral_centroid, ultra_high_energy_ratio, mid_energy_ratio, ultra_high_ratio, low_mid_ratio, damping_ratio, spectral_flatness, resonance_high_ratio, high_freq_slope
Batch 4	env_std, burst_rms, env_mean, resonance_energy_ratio, spectral_bandwidth, temporal_centroid, spectral_centroid, low_mid_ratio, env_kurtosis, spectral_flatness, env_max
Edge v1	spectral_bandwidth, ultra_high_energy_ratio, spectral_centroid, ultra_high_ratio, high_energy_ratio, env_skew, env_kurtosis, spectral_flatness, burst_rms, resonance_skewness, low_mid_ratio

Dataset	Features Improving When Removed
Batch 1	high_freq_slope, spectral_contrast_0, spectral_contrast_1, spectral_contrast_2, spectral_contrast_3, spectral_contrast_4 etc.
Edge v1	resonance_skewness, resonance_q_factor

What information should we focus on?

Saliency Analysis: Analyze which features neural networks consider most important for classification decisions.

Method:

- Train neural networks on each dataset using all 38 features
- Apply gradient-based saliency to rank feature importance
- Identify top features influencing NN predictions

Results - NN Performance and Feature Importance:

Dataset	NN Accuracy	Top 5 Features
Batch 1	95%	spectral_bandwidth, resonance_skewness, spectral_centroid, ultra_high_ratio, zero_crossing_rate
Batch 2	93%	env_min, resonance_high_ratio, damping_ratio, zero_crossing_rate, ultra_high_ratio
Batch 3	97%	spectral_bandwidth, env_decay_rate, spectral_centroid, resonance_q_factor, env_mean
Batch 4	80%	resonance_skewness, spectral_contrast_4, env_std, spectral_centroid, spectral_bandwidth
Edge v1	54%	ultra_high_energy_ratio, ultra_high_ratio, spectral_bandwidth, spectral_flatness, high_energy_ratio

Which Frequencies Matter the Most?

Question: Which frequency ranges contain the most useful information for classifying contact types?

What We Tested:

- 12 different frequency bands (20Hz to 20kHz)
- 5 datasets with different contact scenarios
- Random Forest classifier performance

Which Frequencies Matter the Most?

Band	edge_v1	batch_1	batch_2	batch_3	batch_4
20-200Hz	33.3%	25.0%	25.0%	33.3%	50.0%
200-500Hz	40.0%	90.0%	94.0%	68.7%	63.0%
500-1000Hz	35.2%	70.5%	94.5%	90.7%	66.0%
1000-2000Hz	33.0%	80.5%	94.5%	94.7%	70.0%
2000-4000Hz	38.7%	82.5%	**99.5%**	**99.3%**	69.0%
4000-8000Hz	45.2%	86.5%	**99.0%**	**100.0%**	**78.0%**
8000-20000Hz	**61.4%**	**96.0%**	**99.5%**	**100.0%**	74.0%
20-20000Hz	59.2%	94.5%	100.0%	99.3%	84.0%

Color Coding:

- Green : Top 3 performers per dataset
- Yellow : Full spectrum baseline (20-20000Hz)
- Red : Bottom 3 performers per dataset

Conclusion & Look Ahead

Core Message

We should be able to perform geometrical reconstruction with our setup including two contact scenarios.

Iterative improvement loop:

1. More complex surface / object $\rightarrow$ new data
2. Analyze $\rightarrow$ discover what actually matters
3. Improve finger, signal, or software