Geometric Reconstruction with Acoustic Tactile Sensing

Georg Wolnik

TU Berlin - Biology and Robotics Department

November 22, 2025

Research Question

How can we maximize the information extracted from acoustic signals to achieve reliable contact classification?

Sub-Questions and Corresponding Experiments:

- **Where are the limits for information extraction?** → (1)
- **Which algorithms extract information best?** → (2)
- **Which features contain the most information?** → (3)
- **How do different processing methods compare?** → (4) & (5)

Approach: Systematic analysis of 1,280 audio samples across 5 contact scenarios.

Experimental Datasets

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

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

Five Analysis Methods

1. **Dimensionality Reduction** - PCA and t-SNE visualization
2. **Classification** - Machine learning algorithm comparison
3. **Feature Ablation** - Individual feature importance analysis
4. **Saliency Analysis** - Neural network attention mapping
5. **Impulse Response** - Physical acoustic characterization

Experiment 1: Dimensionality Reduction

Purpose: Visualize class separability in reduced feature space
Method:

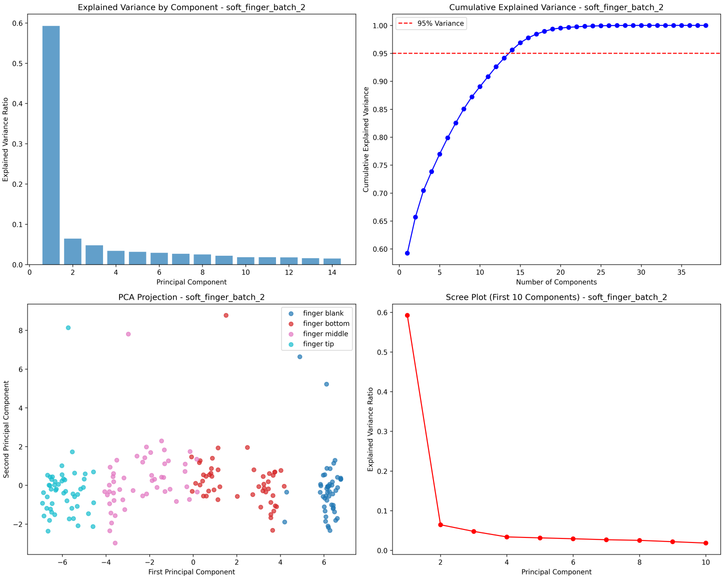
- Principal Component Analysis (PCA)
- t-SNE mapping with silhouette score calculation
- Components needed for 95% variance explained

Results:

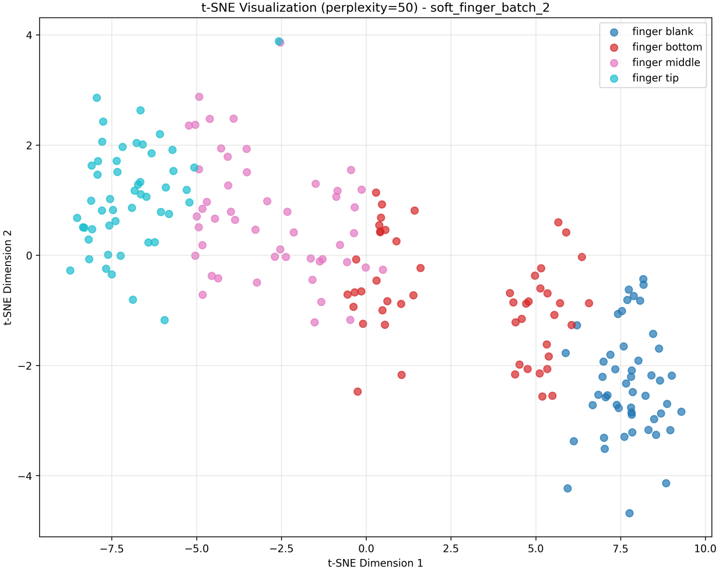
Dataset	Silhouette Score	Components (95%)
Paper clip detection	0.463	12
Edge (Batch 3)	0.347	14
Position (Batch 1)	0.184	16
Position (Batch 2)	0.191	16
Edge (Edge v1)	-0.015	17

Observation: Paper clip detection shows best class separation, while Edge v1 shows class overlap.

PCA Analysis - Batch 2



t-SNE Analysis - Batch 2



Experiment 2: Classification

Purpose: Compare machine learning algorithm performance

Method:

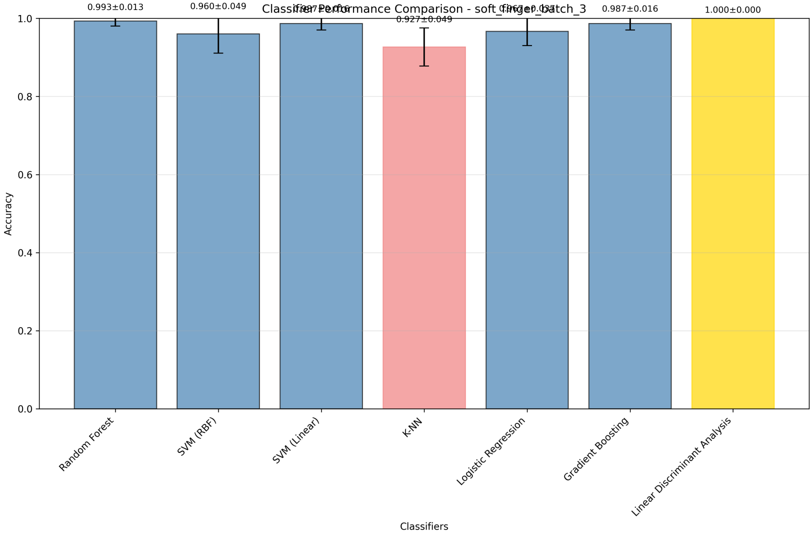
- 7 algorithms tested: Random Forest, SVM, Gradient Boosting, etc.
- 5-fold cross-validation
- Mean accuracy and standard deviation calculated

Results:

Dataset	**Best Algorithm**	**Accuracy (%)**
Edge (Batch 3)	Linear Discriminant Analysis	100.0 ± 0.0
Position (Batch 2)	Random Forest	99.5 ± 1.0
Position (Batch 1)	Random Forest	96.0 ± 3.0
Paper clip detection	SVM RBF	88.0 ± 6.0
Edge (Edge v1)	Gradient Boosting	67.1 ± 2.9

Observation: Performance decreases with task complexity. Edge detection (Batch 3) achieves highest accuracy.

Classifier Performance - Batch 3



Experiment 3: Feature Ablation

Purpose: Determine which acoustic features contribute most to classification

Method:

- Remove each feature individually
- Measure performance drop
- Group features by category (spectral, temporal, statistical, perceptual)

Results - Feature Category Importance:

Dataset	Spectral	Temporal	Statistical	Perceptual
Paper clip detection	67%	8%	12%	18%
Edge Binary	42%	58%	15%	12%
Position	34%	19%	12%	16%
Edge 3-class	28%	22%	20%	19%

Observation: Paper clip detection relies on spectral features, edge detection on temporal features.

Top Features Comparison Across Datasets

Rank	Position Batch 1	Position Batch 2
1	spectral_centroid	high_energy_ratio
2	spectral_bandwidth	ultra_high_energy_ratio
3	spectral_flatness	low_mid_ratio
4	spectral_contrast_0	spectral_centroid
5	spectral_contrast_1	spectral_bandwidth
Drop	0-0.5%	0.5%

Rank	Edge Detection	Paper Clip Detection
1	spectral_centroid	freq_response_spread
2	spectral_bandwidth	high_freq_energy
3	spectral_rolloff	high_freq_damping
4	spectral_flatness	low_energy_ratio
5	spectral_contrast_0	resonance_energy_ratio
Drop	0.7%	2.0%

Key Observations:

- **Spectral features** dominate position and edge detection
- **Impulse response features** critical for paper clip detection
- **Energy ratios** important for position detection (Batch 2)

Experiment 4: Saliency Analysis

Purpose: Analyze which features neural networks consider most important

Method:

- Train neural networks on each dataset
- Apply gradient-based saliency to identify feature importance
- Compare with traditional feature ablation results

Results - Feature Importance Comparison:

Dataset	NN Saliency Top 2	Feature Ablation Top 2
Paper clip detection	freq_response_spread, high_freq_energy	freq_response_spread, high_freq_energy
Position (Batch 1)	spectral_centroid, spectral_bandwidth	spectral_centroid, spectral_bandwidth
Position (Batch 2)	high_energy_ratio, ultra_high_ratio	high_energy_ratio, ultra_high_energy_ratio
Edge detection	spectral_centroid, spectral_bandwidth	spectral_centroid, spectral_bandwidth

Key Finding: Neural networks and traditional methods show strong agreement on feature importance.

Performance Note: NN: 54-95%, Traditional ML: 67-99%

Experiment 5: Impulse Response Analysis

Purpose: Extract system transfer functions independent of input signal characteristics

Method:

- Deconvolution: $H(f) = Y(f)/X(f)$
- Extract 15 impulse response features (resonance, damping, frequency response)
- Combine with 38 acoustic features $\rightarrow$ 53 total features

Key Advantage: Impulse response features characterize physical system properties, independent of input signal design.

Results: Different contact states show distinct acoustic signatures:

- Contact vs Edge vs No Contact: measurable differences in duration, resonance, Q-factor
- 3 of top 6 most important features (saliency analysis) come from impulse response

Summary of Results

Classification Performance Ranking:

1. Edge Detection (Batch 3): 100% accuracy
2. Position Detection (Batch 2): 100% accuracy
3. Position Detection (Batch 1): 96% accuracy
4. Paper Clip Detection: 88% accuracy
5. Edge Detection (Edge v1): 67% accuracy

Key Findings:

- Task complexity correlates with classification difficulty
- Different tasks require different acoustic feature types
- Traditional ML methods generally outperform neural networks
- Physical acoustic differences exist between contact states
- Algorithm choice affects performance significantly

Conclusion

Sub-question Answers:

- *Extraction limits?* Perfect classification achievable (100
- *Best algorithms?* Linear Discriminant Analysis (edges), Random Forest (position), SVM RBF (paper clips), Gradient Boosting (multi-class)
- *Most informative features?* Spectral (paper clips), temporal (edges), impulse response (top 6 features)
- *Processing methods?* Traditional ML outperforms neural networks by 1-13%

Dataset Dependency: Performance varies significantly (67%-100%) across datasets. Continuous validation required with new data batches.

Results reflect performance on specific datasets and require ongoing validation for broader applicability.