2025 IEEE World Haptics Conference (WHC)
Suwon Convention Center, Suwon, Korea
July 8 ~ 11, 2025

CGAN-based Data Augmentation for Improved Tactile Estimation

Yuta Matsumori[1] and Kenjiro Takemura[2]

I. Introduction

Recently, tactile sensation has gained importance in product development for enhancing perceived product value. Accordingly, tactile quantification has progressed to objectively evaluate its impact on product quality in texture. In particular, systems estimating sensory values like "smooth" or "rough" from physically acquired data upon tracing are anticipated for objective understanding. Although machine learning enhances estimation accuracy, it typically necessitates large datasets. The acquisition of tactile data, however, is both resource-intensive and time-consuming.

To address this issue of data acquisition, data augmentation could be one of the solutions. This study proposes to use Conditional Generative Adversarial Networks (CGANs) for tactile data augmentation. While CGANs have been applied to spectrogram generation for tactile displays [1], we are the first to use them for data augmentation in tactile estimation. By generating synthetic spectrograms of pseudo tracing measurements by CGAN-based generator that retains key tactile characteristics from real measurements, this study aims to enhance the tactile estimation performance of the machine-learning-based tactile estimation model.

II. Methodology

We propose a framework for tactile data augmentation using CGANs, aiming to improve tactile estimation accuracy. The proposed framework, illustrated in Fig. 1, consists of two primary components: (a) tactile estimation and (b) spectrogram generation using CGANs.

A. Tactile Estimation Model

The tactile estimation model receives spectrograms of vibration data—acquired from a tactile sensor during sample tracing—as input, and outputs estimated scores corresponding to the evaluation items listed in Table I, which were used in the sensory evaluation experiment. Since the

TABLE I. Evaluation Items

Japanese	English
Dekobokosuru	Uneven
Hukahukasuru	Feathery
Sarasarasuru	Smooth
Hinyarisuru	Cool

[1] Y. Matsumori is with the Graduate School of Science and Technology, Keio University, Yokohama 223-8522, JAPAN. e-mail: yuta.matsumori@keio.jp
[2] K. Takemura is with the Department of Mechanical Engineering, Keio University, Yokohama 223-8522, JAPAN. e-mail: takemura@mech.keio.ac.jp

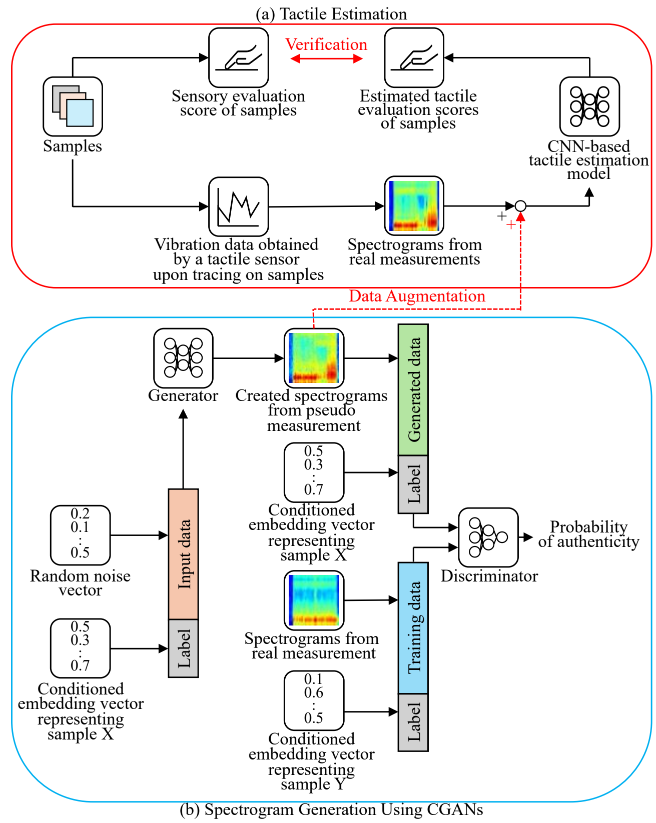

Figure 1. Framework for tactile estimation and tactile data augmentation using CGANs. (a) The tactile estimation model predicts tactile evaluation scores from the spectrograms of vibration data obtained by a tactile sensor upon tracing on samples, potentially enhanced by data augmentation. (b) Generator creates spectrograms of vibration from pseudo tracing measurements based on the embedding vectors representing the sample types and noise, Discriminator evaluates the authenticity of the created spectrograms. The generator and the discriminator are trained adversarially to develop the effective generator.

model utilizes spectrogram images as input, its architecture comprises convolutional layers for feature extraction, followed by a flattening layer and fully connected layers for tactile score estimation, i.e., CNN. Spectrograms were employed as they effectively capture time-frequency characteristics of vibration signals related to tactile perception, while being compatible with both CNN and GANs. The model was trained on a dataset (cf. II.D) using mean square error (MSE) loss as the loss function with an 80:20 split for training and validation.

B. Data Augmentation using CGANs

The CGANs consist of a generator and a discriminator that compete against each other during training to improve performance. The generator transforms conditioned embedding vector, representing the sample types, and random noise into pseudo-spectrograms using transposed convolutional layers. The discriminator evaluates the authenticity of spectrograms using convolutional layers, conditioned on the embedding vectors representing sample types. For each sample, the variance of 10 pseudo-spectrograms and that of 10 real spectrograms were calculated. The sum of absolute errors between these variances across samples was used to select the best generator, ensuring data diversity, which is crucial for enhancing the generalization capability of the tactile estimation model during training.

C. *Dataset Preparation*

Dataset was employed from the previous study of the authors [2]. Vibration data was collected using a piezoelectric sensor upon tracing 15 samples at constant speed of 20 mm/s and normal force of 0.5 N, which were determined based on human tactile perception behavior. These signals were converted into spectrograms via Short-Time Fourier Transform. Forty participants (20M/20F; mean age 22.6±0.98) evaluated each sample using the Semantic Differential (SD) method on a 7-point scale using four evaluation items (Table I). The experiment protocol was approved in advance by the Bioethics Board of the Faculty of Science and Technology, Keio University (2022-001). The mean evaluation score across participants served as ground truth. The samples were clustered into three classes based on the evaluation scores, with one test sample selected from each class. These three test samples—Vinyl, Rubber, and Wood—were individually excluded from both the CGANs and the tactile estimation model training to serve as unknown samples for the estimation verification.

D. *Evaluation of Tactile Estimation Model trained w/ or w/o data augmentation*

To evaluate the effectiveness of the proposed data augmentation, we prepared two datasets for training the tactile estimation model as Table II, and compared their estimation errors for test samples.

III. RESULT

A. *Quality of Generated Spectrograms*

Fig. 2 compares a real spectrogram and a CGAN-generated spectrogram of a leather sample. The t-SNE visualization in Fig. 3 shows the distribution of real and generated spectrograms in a feature space for all samples except for the Vinyl sample, which was used as the test sample. The generated data cluster near the real data, although some distributional bias is observed; nevertheless, the results suggest that effective data augmentation can be expected.

TABLE II. TWO DATASETS

Dataset	Real data	Pseudo data	Total
Real data only	10×14 samples	0	140
Augmented dataset	10×14 samples	90×14 samples	1400

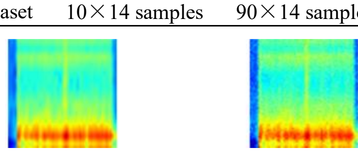

Figure 2. Comparison of real (left) and CGAN-generated (right) spectrograms from a leather sample.

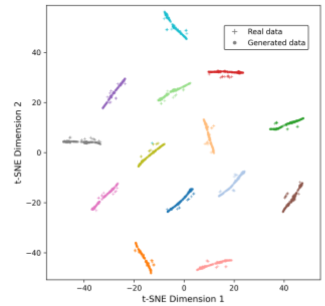

Figure 3. t-SNE visualization of real (cross marks) and generated (dots) spectrogram distributions in feature space, with each color representing a different sample type.

TABLE III. MEAN ABSOLUTE ERRORS ON THE VINYL TEST SAMPLE

Evaluation Items	Real data only	Augmented dataset	Improvement ratio
Dekobokosuru	1.026	1.042	-1.6 %
Hukahukasuru	1.703	1.304	23.4 %
Sarasarasuru	0.491	0.489	0.4 %
Hinyarisuru	0.566	0.497	12.2 %
Total	3.786	3.332	12.0 %

B. *Improvement in Tactile Estimations*

Table III summarizes the mean absolute errors between predicted and actual sensory evaluation scores for each evaluation item when the Vinyl sample is treated as unknown. The models trained with augmented dataset showed a tendency toward lower absolute prediction errors compared to those trained with real data only. Similar improvements were observed in the other two test samples, with total mean absolute errors decreasing by 6.9% for the Rubber sample and 12.5% for the Wood sample.

C. *Discussion*

A consistent reduction in total mean absolute error across all test samples suggests that the CGAN-generated data contributed positively to the estimation model performance. This indicates that the generated spectrograms captured structural characteristics of the real data, allowing them to function effectively as supplementary training inputs.

However, the improvement varied depending on the evaluation items, and in some cases, the estimation accuracy decreased. t-SNE visualizations showed that while the generated data were generally located near the real data clusters, distributional bias remained. This bias may have limited the effectiveness of the data augmentation.

IV. CONCLUSION AND FUTURE WORK

This study demonstrated that CGAN-generated pseudo-spectrograms can contribute to improving the tactile estimation performance. Future work will focus on addressing distributional bias in generated data to achieve more diverse and meaningful data augmentation.

REFERENCES

[1] Y. Ujitoko, Y. Ban, and K. Hirota, “GAN-based fine-tuning of vibrotactile signals to render material surfaces,” *IEEE Access*, vol. 8, pp. 16656–16661, 2020.

[2] M. Sagara, and K. Takemura, “Texture classification model based on temporal changes in vibration using wavelet transform,” *2022 IEEE Sensors*, pp.01–04, 2022.