



OPEN ACCESS

EDITED BY

Rocco Antonio Romeo,
Italian Institute of Technology (IIT), Italy

REVIEWED BY

Barış Can Yalçın,
University of Luxembourg, Luxembourg
Siao Wang,
Xtreme Intelligence LLC, China

*CORRESPONDENCE

Deepesh Kumar,
✉ deepesh.bme@iitbhu.ac.in

[†]These authors have contributed equally to
this work and share first authorship

RECEIVED 29 December 2024

ACCEPTED 02 June 2025

PUBLISHED 26 June 2025

CITATION

Patel N, Rana R, Kumar D and Thakor NV
(2025) SuperTac - tactile data
super-resolution via dimensionality reduction.
Front. Robot. AI 12:1552922.
doi: 10.3389/frobt.2025.1552922

COPYRIGHT

© 2025 Patel, Rana, Kumar and Thakor. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

SuperTac - tactile data super-resolution via dimensionality reduction

Neel Patel^{1†}, Rwik Rana^{1†}, Deepesh Kumar^{2*} and
Nitish V. Thakor³

¹Department of Mechanical Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar,
Gujarat, India, ²School of Biomedical Engineering, Indian Institute of Technology (BHU) Varanasi,
Varanasi, Uttar Pradesh, India, ³Department of Biomedical Engineering, Johns Hopkins University,
Baltimore, MD, United States

The advancement of tactile sensing in robotics and prosthetics is constrained by the trade-off between spatial and temporal resolution in artificial tactile sensors. To address this limitation, we propose SuperTac, a novel tactile super-resolution framework that enhances tactile perception beyond the sensor's inherent resolution. Unlike existing approaches, SuperTac combines dimensionality reduction and advanced upsampling to deliver high-resolution tactile information without compromising the performance. Drawing inspiration from the spatiotemporal processing of mechanoreceptors in human tactile systems, SuperTac bridges the gap between sensor limitations and practical applications. In this study, an in-house-built active robotic finger system equipped with a 4 × 4 tactile sensor array was used to palpate textured surfaces. The system, comprising a tactile sensor array mounted on a spring-loaded robotic finger connected to a 3D printer nozzle for precise spatial control, generated spatiotemporal tactile maps. These maps were processed by SuperTac, which integrates a Variational Autoencoder for dimensionality reduction and Residual-In-Residual Blocks (RIRB) for high-quality upsampling. The framework produces super-resolved tactile images (16 × 16), achieving a fourfold improvement in spatial resolution while maintaining computational efficiency for real-time use. Experimental results demonstrate that texture classification accuracy improves by 17% when using super-resolved tactile data compared to raw sensor data. This significant enhancement in classification accuracy highlights the potential of SuperTac for applications in robotic manipulation, object recognition, and haptic exploration. By enabling robots to perceive and interpret high-resolution tactile data, SuperTac marks a step toward bridging the gap between human and robotic tactile capabilities, advancing robotic perception in real-world scenarios.

KEYWORDS

mechanoreceptors, robotic finger, tactile sensor, tactile super-resolution, texture, variational autoencoder

1 Introduction

In traditional robotics, vision has been the primary sensory modality. However, as robots are increasingly deployed in unstructured environments and tasked with complex object manipulation, the sense of touch becomes indispensable. Tactile sensing

is a critical component of robotic or prosthetic perception, enabling machines to interact with their surroundings through touch. It plays a pivotal role in applications such as object recognition, material classification, robotic manipulation, and haptic exploration (Pyo et al., 2021; Meribout et al., 2024). A robust sense of touch allows robots to perform tasks that require precise interaction, such as handling fragile objects or exploring intricate surfaces. Despite significant advancements, artificial tactile sensing systems remain limited in achieving the high resolution and efficiency of human touch perception. A key challenge is the trade-off between spatial and temporal resolution in tactile sensors. Increasing spatial resolution often reduces temporal resolution due to hardware constraints such as sampling rates and communication bandwidths (Wang et al., 2023; Huang et al., 2025).

The human tactile system provides an exemplary model for addressing this trade-off. Mechanoreceptors in human skin process tactile information with high spatial and temporal resolution, and the brain integrates this data to enable rapid and precise tactile perception (Johansson and Flanagan, 2009). Remarkably, the human tactile system achieves hyperacuity, distinguishing tactile stimuli separated by as little as 0.3 mm (Abraira and Ginty, 2013), finer than the receptive field of any single mechanoreceptor (approximately 2 mm). This capability arises from the population-level encoding of tactile information, where the spatiotemporal patterns of mechanoreceptor activation across neighborhoods are processed synergistically. This biological inspiration underscores the need for artificial systems capable of similar spatiotemporal processing, bridging the gap between sensor limitations and practical applications.

Previous research aimed at enhancing tactile resolution has predominantly concentrated on developing specialized sensor architectures and advanced fabrication techniques (Wang et al., 2023). For example, Lu et al. (2024) designed a biomimetic soft tactile sensor inspired by the Pacinian corpuscle, optimizing the soft silicone layer for super-resolution. Li et al. (2022) utilized a high-throughput laser manufacturing method to achieve fine spatial resolution (0.7 mm) with minimal crosstalk. Yan Y. et al. (2021) developed a flexible, self-powered triboelectric sensing array via laser direct writing on laser-induced graphene, enabling high-resolution (8 dpi) real-time sensing. Similarly, Zhang et al. (2022) introduced a fast-photocurable solid-state conductive ionoelastomer (SCIE) that supports high-resolution 3D printing of robust, stretchable tactile sensors. While these approaches have demonstrated impressive performance, they are inherently tied to specific sensor designs and fabrication processes. While these fabrication-centric strategies have advanced tactile sensor performance, they are often resource-intensive, sensor-specific, and limited in scalability. As tactile sensing applications expand across diverse platforms, there is a growing need for algorithmic approaches to super-resolution that can enhance tactile resolution independent of the underlying sensor hardware.

More recently, a new wave of tactile sensing strategies has focused on integrating various tactile super-resolution algorithms with existing or custom-made tactile sensors to address the limitation of traditional tactile sensors sensing (Yu and Liu, 2025). Few studies mimicked the tactile sensing and encoding strategy used by human mechanoreceptors (neuromorphic tactile sensing) to achieve high-resolution tactile sensing and showed

applications in texture classification, edge detection and slip detection (Kumar et al., 2020; Sankar et al., 2019; Parvizi-Fard et al., 2021). Other studies used various machine learning based approach for tactile super-resolution. For instance, a deep neural network-based reconstruction framework, EIT-NN, was proposed to enhance the performance of electrical impedance tomography (EIT)-based sensors, improving spatial resolution and sensitivity while maintaining simplicity in sensor design (Park et al., 2021). Similarly, the Local Message Passing Network (LoMP) enabled high-resolution calibration of piezoresistive sensor arrays using limited single-touch data, addressing calibration challenges in multi-touch scenarios (Kim et al., 2021). Another significant development involves the use of soft magnetic skin for tactile sensing, which decouples normal and shear forces and achieves super-resolution through deep learning algorithms (Yan Z. et al., 2021). Another approach by Wu et al. (2022) introduces TactileSRCNN and TactileSRGAN, which adapt image super-resolution techniques such as CNNs and GANs to upscale low-resolution tactile patterns from taxel-based sensors by a factor of 100, enabling multi-point contact detection from a single tap. Another method (Oller et al., 2023) focuses on modeling the dynamics of deformable tactile membranes by combining 3D geometric data and proprioceptive feedback to predict sensor deformation and improve manipulation control. Similarly, (Ouyang et al., 2024) presents a high-resolution piezoresistive sensor array integrated with machine learning algorithms, achieving fine spatial and temporal resolution and demonstrating 98.9% accuracy in shape recognition. These studies collectively highlight the growing role of machine learning in achieving tactile super-resolution and precise pattern recognition. These advancements have enabled precise tactile feedback for tasks such as adaptive grasping and teleoperation, emphasizing the potential of combining advanced sensor designs with computational models.

Despite these innovations, a generalized approach to achieving tactile super-resolution using standard low-resolution sensors remains a challenging task. Existing methods often require specialized hardware or extensive calibration, limiting their scalability. To address these limitations, we propose SuperTac, a real-time, hardware-efficient tactile super-resolution framework that operates using standard low-resolution tactile sensors. By combining dimensionality reduction using a Variational Autoencoder (VAE) and advanced upsampling using Residual-In-Residual Blocks (RIRB), our method delivers fourfold spatial resolution improvement while maintaining over 50 frames per second (FPS) throughput. This makes SuperTac a promising step toward a generalizable framework for real-time tactile super-resolution, offering high performance without relying on complex hardware or visual sensors.

To validate the SuperTac framework, we designed an experimental setup featuring an in-house-built active robotic finger system equipped with a 4×4 tactile sensor array. This system palpated textured surfaces to generate spatiotemporal tactile maps as input data for the SuperTac network. We experimented with super-resolution outputs of 8×8 , 16×16 and 32×32 . Among these, the 16×16 resolution offered the best trade-off between the image detail and reconstruction quality, leading to significantly improved tactile representation. The results demonstrate a fourfold

improvement in spatial resolution, producing 16×16 super-resolved (SR) tactile images.

While this study uses a tactile sensor similar in design to the one employed by Kumar et al. (2020), our objective and methodological approach are fundamentally different. Kumar et al. focused on encoding spatiotemporal features from low-resolution tactile data for classification tasks using standard neural networks. In contrast, our work introduces *SuperTac*, a novel framework that integrates dimensionality reduction via a Variational Autoencoder with Residual-In-Residual Blocks to achieve real-time super-resolution of tactile data. This enables multifold increase in spatial resolution, leading to significantly improved texture classification performance and enabling broader applications in high-resolution tactile perception. Additionally, SuperTac is designed to be sensor-agnostic and operates in real time, offering scalability that was not addressed in Kumar et al.'s work.

The major contributions of this work include.

1. Development of a novel tactile super-resolution framework that integrates dimensionality reduction and advanced upsampling.
2. Demonstration of a fourfold improvement in tactile sensor resolution with computational efficiency suitable for online deployment.
3. Validation of the framework through an active robotic finger system for tactile data collection.
4. Significant improvement in texture classification accuracy using super-resolved tactile data.

This paper is organized as follows: Section 2 describes the materials and methods used in this study. Section 3 details the SuperTac framework, including its architectural components and training methodology. Section 4 presents the results and Section 5 discusses the findings. Finally, Section 5 concludes the paper with insights and future directions.

2 Materials and methods

2.1 Tactile sensor array

This study uses a fabric-based piezoresistive tactile sensor for the robotic palpation experiment. We use a similar tactile sensor previously used by Kumar et al. (2020). It is a 2D array of 16 tactile sensing elements (taxels) arranged in a 4×4 grid, within an area of 13×13 mm (Figure 1). A piezoresistive cloth is sandwiched between conductive traces arranged as rows and columns. The width of traces is 2 mm with a 1 mm spacing between consecutive traces, and therefore each taxel has a size of $2 \text{ mm} \times 2 \text{ mm}$. The tactile data were recorded at the sampling rate of 300 Hz per taxel.

2.2 Tactile stimuli and robotic palpation

2.2.1 Design of tactile stimuli

We used two categories of graded textures for our experiments, i.e., ridges and bumps. These textures were fabricated using a 3D printer and PLA plastic as the printing material. As shown in Figure 2, ridges are triangular protrusions, and bumps

are semi-circular protrusions. The three textures of each type have a 12 mm, 6 mm, and 4 mm distance between each protrusion. Varying the distance between subsequent protrusions allows the dataset to be diverse, and the proposed network would learn to generate output as a convex combination of the known bumps and ridges.

2.2.2 Design of robotic finger

The tactile sensor is integrated into an in-house-built robotic finger system (Figure 3a). The finger consists of three components: the main body, mid-piece, and tactile fingertip (Figure 3b). A helical spring is incorporated into the design to function as a suspension system, enabling the finger to bend passively in response to external forces. This passive compliance allows the finger to conform smoothly to various terrains and surfaces. The stiffness of the spring ensures that the bending occurs in a controlled manner, providing stability during operation. To achieve precise spatial movement, the main body of the robotic finger is securely mounted onto the nozzle of a 3D printer.

2.3 Experimental protocol

The entire experimental setup and data acquisition system has 4 phases. The phases are categorized according to the rectangular movement of the finger to palpate over textured plates. Figure 4 shows the four phases; onset, sliding, release, and re-position.

For each texture plate, eight sliding actions are performed to cover the width of the textured plate. Two trials were performed for each of eight sliding actions for each texture. The texture plates were fixed on the 3D printer base and the robotic finger palpated over the textured plates during the palpation. The tactile data were collected for sliding speeds at 5 mm/s and 10 mm/s.

2.4 Super-Tac algorithm

The proposed approach for tactile data super-resolution stems from the idea that points on an image depict their positions with respect to each other, and extracting these features helps capture information about the point and its surrounding. Along with the spatial component, the temporal features give us an idea about the context of the data in the current time step and use information from preceding time steps to extract underlying features. Combining both spatial and temporal feature extractors enable the extraction of rich features from the input sequence (Kumar et al., 2020). To capture the temporal features, a vast percentage of methods employ a sliding frames window along with neural network architectures like bidirectional recurrent convolutional networks (BRCN) (Jo et al., 2018) and long short-term memory networks (LSTM) (Huang et al., 2015). Our methodology combines the aspects of tactile sensing with image processing to generate super-resolution data. We have visualized tactile data in the form of images and used deep learning-based algorithms to achieve tactile super-resolution. We take inspiration from VAE to capture the latent features, which are simpler to process, and we modify the residual blocks presented in Zhang et al. (2018) as our upsampler. Thus, we intend to reduce the dimensions of the input data and then use the dimensionally reduced latent parameters to fabricate the SR image.

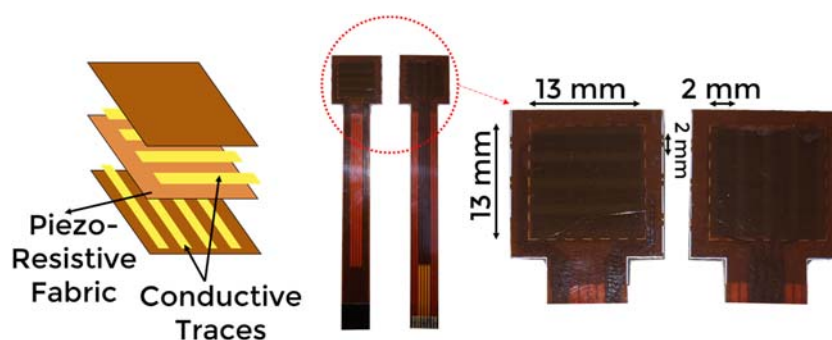


FIGURE 1
An exploded view of the tactile sensor array used in the experiment.

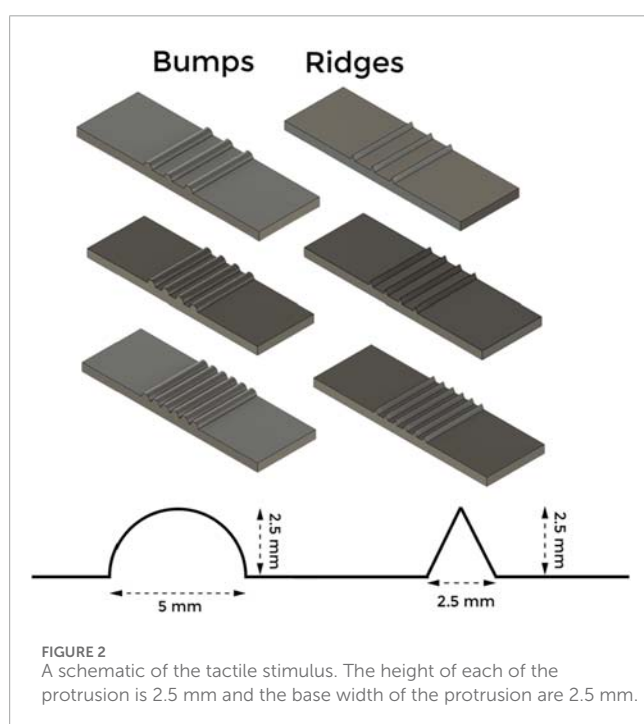


FIGURE 2
A schematic of the tactile stimulus. The height of each of the protrusion is 2.5 mm and the base width of the protrusion are 2.5 mm.

2.4.1 Data visualization and pre-processing

The tactile sensor used in this study has 16 tactile sensing elements stacked in a grid manner, and therefore we visualize the data in 2 forms. The 16 channels are flattened along with the columns and rows, denoting the finger's movement in the direction of columns and rows, respectively. "N" such flattened row-vectors are stacked over each other; N is the number of timesteps (Ts). Each row in the visualization corresponds to a particular timestep, and each cell in a row corresponds to tactile data from the 16 taxels.

The movement of the 3D printer nozzle is constrained to the horizontal plane only. The tactile sensor provided the output data in the range of 0 V–2.5 V, which was linearly mapped to a scale of [0–255], corresponding to an 8-bit resolution commonly used in image processing. We try to visualize the obtained tactile data in the form of 4×4 images, enabling us to apply the concepts from image processing in our analysis.

2.4.2 Variational autoencoders

A Variational Autoencoder (VAE) (Kingma and Welling, 2014) provides a probabilistic manner for describing an observation in latent space. It has two main components; the encoder and the decoder. The latent parameters are sampled from the normal distribution using the mean and variance of the output of the encoder network. VAE is suited for dimensionality reduction of tactile sensors into latent features. Biological neurons pass the latent information in the form of neuronal spikes. Sorting the neuronal spikes based on a gaussian mixture model (GMM) (Souza et al., 2019) is well known and practiced. This suggests that the latent features can be represented by Gaussian curves or sampled from normal distributions. This is what VAE exactly does; proper training of VAE using Kullback–Leibler divergence Loss (KLD) loss ensures that latent features are sampled from the standard normal distribution (a particular case of a gaussian curve). Thus, VAE's functionality of representing information is similar to the way actual biological neurons do and thus is a good fit for our network.

$$\left. \begin{aligned} x &= x_1, x_2, x_3, x_4 \\ f(x_1) + h_1(x_i) &= \mu_i, f(x_i) + h_2(x_i) = \sigma_i \\ l_{r_i} &\sim N(\mu_i, \sigma_i) \Rightarrow l_{r_i} = \mu_i + \epsilon_i \sigma_i \\ LR &= [l_{r_1}, l_{r_2}, l_{r_3}, l_{r_4}] \\ Output &= [g(l_{r_1}), g(l_{r_2}), g(l_{r_3}), g(l_{r_4})] \end{aligned} \right\} \quad (1)$$

The Equation 1 is a simplified mathematical model for our context, f is the encoder network with input from the VAE stack $x_i | i \in \{1, 2, 3, 4\}$, and h_1 and h_2 are identical but disjoint functions added to the final layer of the encoder function, whose weights and gradients are calculated separately to obtain the mean and variance. We sample from a normal distribution to obtain the latent representation l_{r_i} . Finally, g is the decoder network that takes the latent data as input.

We hypothesize that the input image of the tactile data can be reduced to latent representations. Using VAE, the $N \times 16$ input data can be expressed by $n \times 4$ data array, where N is the total time steps (temporal component) of the input sequence, and n is the dimensions of the latent parameter. The 16 columns are separated into groups of four based on horizontal visualization

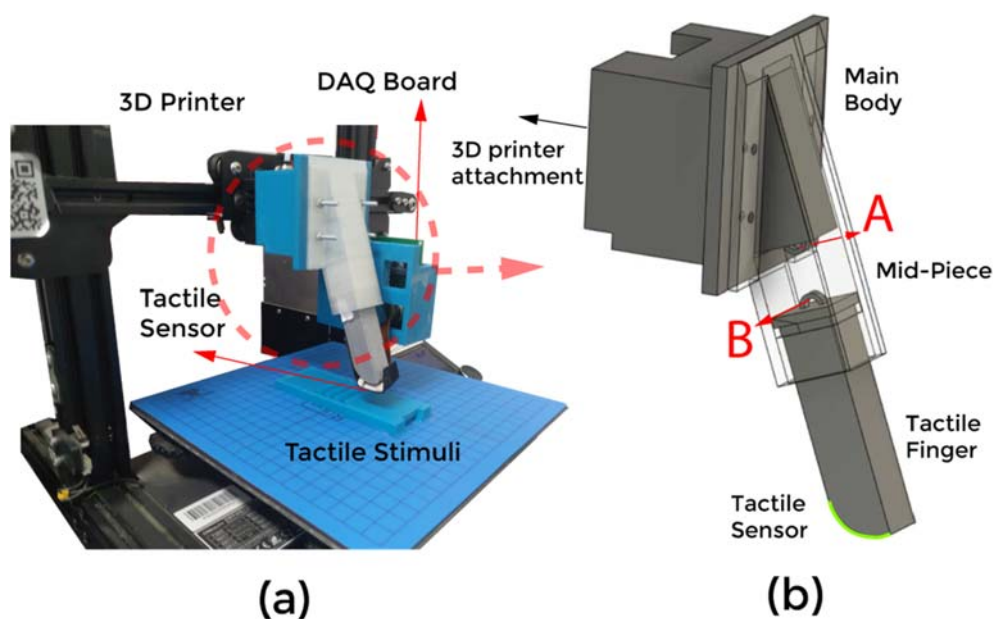


FIGURE 3
(a) The experimental setup with the finger attached to the 3D printer. **(b)** An enlarged view of the finger containing three parts; Main body, Mid-Piece and the finger. The suspension system is attached to points (A, B).

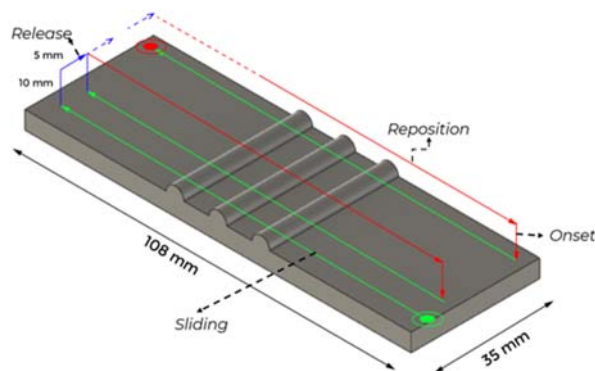


FIGURE 4
 The tactile finger traverses on the tactile stimuli in 4 phases; Onset: Positioning the finger above the initial point. Sliding: Palpating the finger over the texture stimuli. Release: Breaking the contact of the finger and the stimuli and offsetting the finger for the next sliding action. Reposition: Bringing the finger to the next initial point which is at a horizontal offset to the previous initial point.

or vertical visualization (Figure 5). Each of the $N \times 4$ data is down-sampled to an $n \times 1$ array using a separate VAE network. We concatenate the down-sampled arrays to get an encoding with the shape $n \times 4$. The temporal nature of the data (as discussed in Section 2.4.1) ensures that the VAE extracts spatio-temporal features.

We chose a VAE for dimensionality reduction because it learns a smooth, structured latent space that helps generalize tactile features. While VAEs can sometimes produce slightly blurrier outputs

than standard autoencoders, we found the trade-off acceptable for our task.

2.4.3 Residual upsampler

A study by Cheng et al. (2019) introduced a method to effectively use an encoder-decoder network with residual skip connections to obtain SR images from input blurred images. The up-sampling block contains the Residual In Residual Block (RIRB), which combines the data from the previous layers and concatenates it with the processed data. The input to the first RIRB block is the latent features from the encoder network. Subsequent RIRB blocks are connected to each other in a sequential manner. RIRB contains several Residual Channel-wise Attention Blocks (RCAB) (Zhang et al., 2018), a convolution layer, and a skip connection. RCAB adopts channel-wise attention mechanism to adaptively distinguish the significance of the channels of the input data to the RCAB Block. It gives more significance to some of the extracted features among the input channels. In our case, given that we are dealing with latent features of the tactile data, not all the features are equally important. RCAB helps in giving significance to certain features, unlike commonly used residual blocks wherein the features are treated fairly. In our implementation, we use similar blocks but use them as our up-sampler block (by adding a transpose convolution layer) to selectively extract features and get SR output. Thus, the RIRB block in our case works both as a selective feature extractor and a upsampler to decode the latent features to super-resolution outputs.

For upsampling, we used RIRB, which are effective at preserving fine details in super-resolution. Though alternatives like sub-pixel convolution or UNet could also work, the proposed setup is chosen to achieve good balance between performance and efficiency for real-time use.

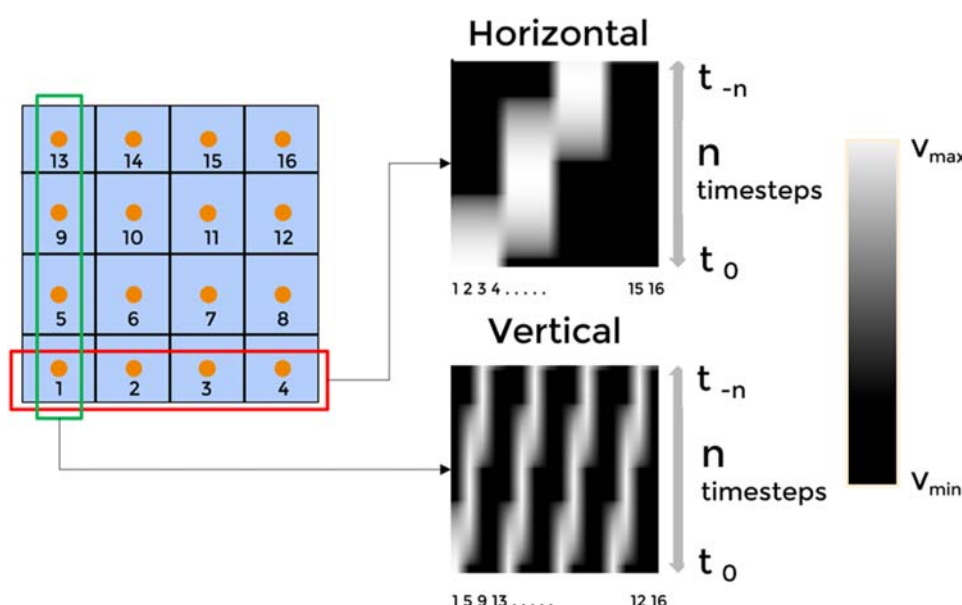


FIGURE 5

The tactile sensor comprises 16 taxels, numbered from 1 to 16. The data from these taxels can be flattened and vectorized either along rows or columns. Time series corresponding to each row or column are then stacked over time to form a spatio-temporal representation. The right panel illustrates both horizontal and vertical vectorization strategies, where each row or column's temporal dynamics are encoded as image-like patterns.

2.5 Synthetic dataset generation

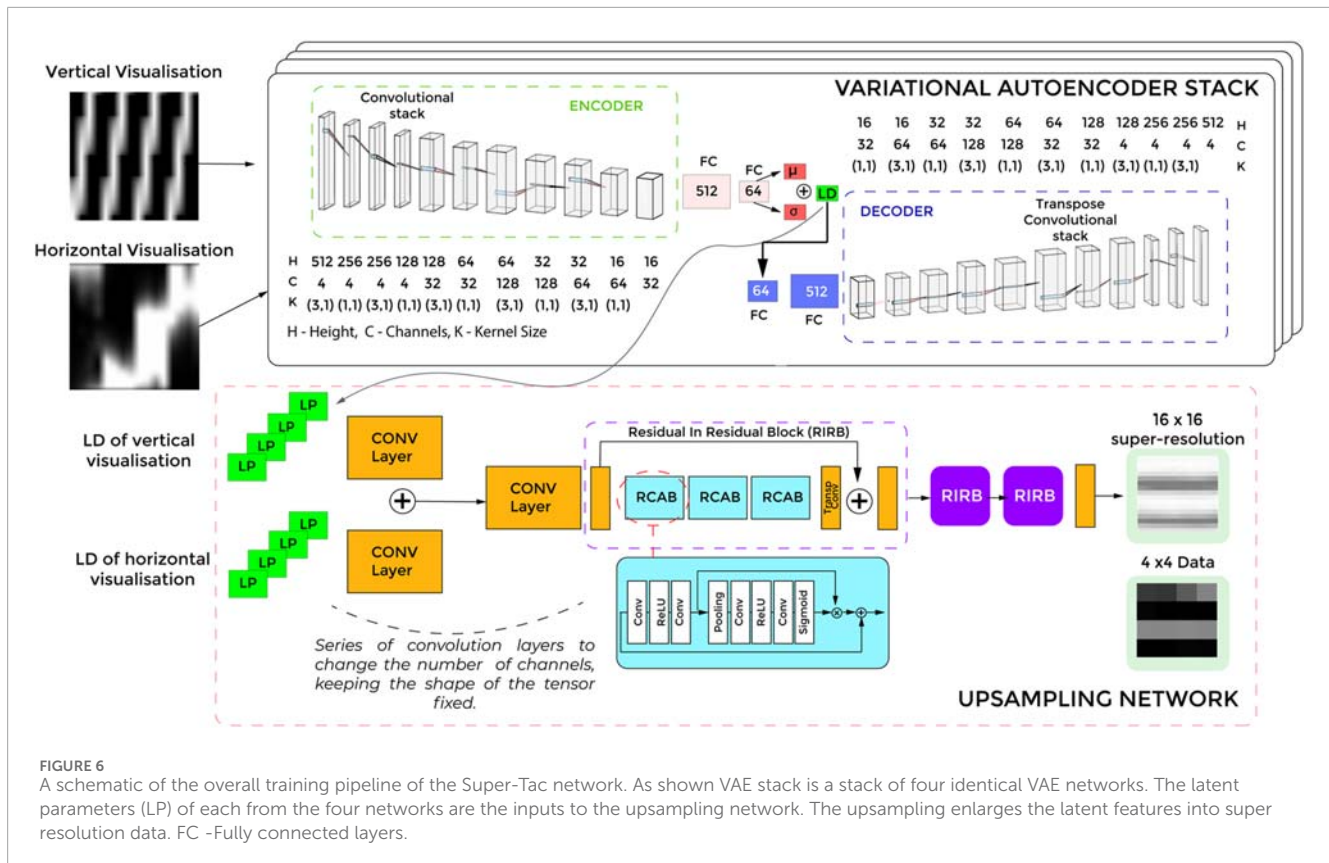
As mentioned earlier, we converted the raw tactile data obtained from robotic palpation over textured surfaces into a 4×4 tactile images by linearly interpolating the voltage data, from the tactile sensor to grayscale. This 4×4 images are then passed through the SuperTac network to obtain 16×16 super-resolved tactile images. However, to assess the quality of obtained 16×16 tactile images, we needed an ideal representation of the textured surface in the form of 16×16 image within the same $13 \text{ mm} \times 13 \text{ mm}$ of tactile sensing area. Therefore, we created synthetic dataset of 16×16 image for each timestep. Knowing the speed of the of the finger palpation and geometry of the texture enabled us in generating the ideal 16×16 tactile image, with each pixel denoting $1/16$ th of the $3 \text{ mm} \times 3 \text{ mm}$ tactile sensor. For each time-step, the location of the finger on the textured plate is known and thus accordingly the location of the centre points of the 256 pixels (16×16) is calculated and the height of the texture at the any position is taken from the 3D design file of the texture. The height is then linearly-interpolated to grayscale. Also, to make the ground truth more realistic, we have added certain noises like mechanical vibration noise as minor sensor or object movements can create wavy distortions in the readings. Also, pressure can also be inconsistent at some points while palpating over textures. Hence, mechanical vibration noise will incorporate those factors in the ground truth dataset. Further, we added thermal drift, EMI noise, quantization artifacts, and sensor crosstalk noises. It is to be noted that the time interval between subsequent data is taken to be constant i.e., $1/300 \text{ s}$.

3 Training pipeline

The whole training pipeline, as shown in Figure 6 consists mainly of two parts, the VAE Stack and the Upsampling Network. We use an end-to-end approach to train the encodings and the super-resolution networks together. For the VAE stack, we use Kullback–Leibler divergence Loss (KLD Loss) and L2 loss for each VAE network of the stack. L2 Loss penalizes the model based on the Cartesian distance between the input and decoded images. KL divergence term in the loss function makes the distribution of the encoder output as close as possible to the standard multivariate normal distribution.

For the upsampling stack, we use L2 Loss and SSIM (Structural Similarity Index). L2 loss takes care of the Cartesian distance between the super-resolution output and the high-resolution ground truth data. The SSIM is a perceptual image measure, widely used for measuring performance of super-resolution networks. The overall loss of the network is a convex combination of losses from VAE stack and Upsampling network.

The Optimizer used in the process is the Adam optimizer. Our dataset consists of two main textures with a varying number of protrusions; bumps and ridges. The distances between the protrusions have been varied, as stated in Section 2.2.1. The dataset is split evenly between bumps and ridges. The tactile sensor has a sampling rate of 300 samples per second. We collected data over the 6 tactile stimuli (Figure 2) over two speeds of the robotic palpation; 5 mm/s and 10 mm/s . For each stimulus and speed pair, we performed eight sliding actions. This yielded us approximately 466000 numbers of 4×4 images.



4 Results

Figure 7 shows output of Super-Tac network. The input to the network is the tactile data of shape 4×4 (first column). The low-resolution data is upsampled to obtain super-resolution outputs denoted by the final column. Each of the output has a size of 16×16 . The middle column is the simulated ground truth data. The texture plate used in this case consists of both bumps and ridges at varying distances from each-other.

In order to quantify our results, we use SSIM (Structural Similarity Index), PSNR (Peak Signal to Noise Ratio), and Phase Correlation as our metrics. These metrics are used to measure the performance of the SR tactile images to the simulated ideal high resolution tactile images. We also report the frames per second (FPS) of output SR images for the suitability of real-time usage of the proposed SuperTac algorithm.

4.1 SSIM and PSNR

Calculating the SSIM (Wang et al., 2004) and PSNR scores as a metric to determine the performance of super-resolution is a common practice (Wang et al., 2018; Ledig et al., 2017). Since our method involves visualizing the tactile data in an image-like format, we use the SSIM and PSNR metrics for our analysis.

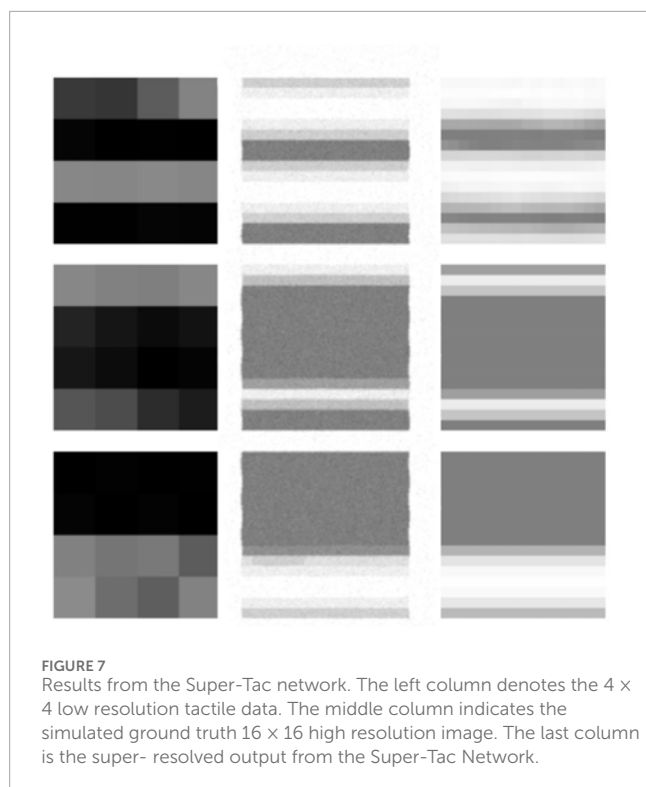
$$PSNR = 20 \log_{10}(MAX) - 10 \log_{10}(MSE) \quad (2)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + 1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3)$$

In Equation 2, the MAX term represents the maximum possible value of each unit cell, while MSE denotes the mean squared error between the reconstructed and reference images. In Equation 3, the μ terms correspond to luminance (mean pixel intensity), and the σ terms capture contrast (standard deviation of pixel values). The constants C_1 and C_2 are introduced to ensure numerical stability, particularly when the luminance or contrast values approach zero. We plot (Figure 8) the average PSNR and SSIM by varying the latent dimensions. Each plotline in the graph is drawn, keeping the number of timesteps constant. We can see that both SSIM and PSNR score gradually ingresses. For the instant when the number of timesteps is 256, the SSIM increased from 0.749 to 0.856 Gradual increase in the SSIM score is due to larger, better, and richer encoding of the input image.

4.2 Frames per seconds (FPS)

The FPS is calculated by considering the data input ($n \times 16$ image) as a frame at a particular time step. We calculate the FPS for all the cases by varying the number of latent dimensions (LD) and the number of timesteps to the network. We observe a gradual but minor increase in the inference speed with a decrease in the number of timesteps (Ts). Super-Tac processes all the data points in the range of 47.1–53.2 FPS (47.1 FPS for 512 Ts and 30 LD and 53.2 FPS for 30Ts and 3 LD), which suggests that the SR image of the input data



can be generated in real-time given the FPS achieved are greater than 20 FPS. Therefore, we get the super-resolution output of the input image in real-time and can thus be used as a live feed when the network is employed on a prosthetic finger or a manipulator hand of an exploration robot. This also gives us almost 20 FPS overhead and thus allows further processing of the super-resolution output if needed and still makes the process real-time. FPS decreases, and PSNR values increase with the increase in latent dimensions, as shown in Table 1 and Figure 5, thus creating a tradeoff of speed v/s performance. This prompts us to consider the number of latent dimensions to be 10 for further analysis. This ensures that we get inference speeds almost close to the case of three latent dimensions but with performance closer to the case of 30 latent dimensions.

4.3 Phase correlation

In some cases, despite achieving accurate inferencing, the PSNR and SSIM values remain low due to the construction of the SR tactile images and ground truth tactile images using binned data samples over fixed time periods. While generating the ground truth, we assume that the time interval between consecutive data points is constant. However, during palpation, there is an inherent margin of error in these intervals, causing them to vary (Figure 9). This variability introduces cumulative errors during data collection. Since SR images are generated from the collected data, which includes this time interval variability, whereas the ground truth relies on an idealized constant interval assumption, the PSNR and SSIM values are influenced by the duration of the binning period. This issue led us to adopt phase correlation as an alternative metric, as it is more robust to shifts between similar images.

We trained separate networks with varying timesteps: 512, 256, 128, 64, and 32, and analyzed the phase correlation (PC) values, which are presented in Table 1. We can see a bell-shaped relationship between timesteps and phase correlation in Table 1, with the highest performance observed at 256 timesteps (94.325 PC). This optimal performance strikes a balance between data resolution and smoothing, leading to the best accuracy in tactile data interpretation.

The above discussed results were obtained using a fourfold super-resolution (16×16) of the 4×4 input tactile image. To evaluate the effect of varying the super-resolution factor, we extended the analysis to include twofold (8×8) and eightfold (32×32) super-resolutions by modifying the SuperTac network accordingly. Figure 10 presents the qualitative results for these output resolutions. We noticed that 8×8 super-resolved output has less details i.e., it is limited in spatial resolution and 32×32 super-resolved image can introduce more artifacts, as the upscaling factor increases, small errors in estimation or interpolation get amplified, leading to artifacts like blurring, ringing or unrealistic patterns. Quantitative evaluation using SSIM and PSNR metrics, computed over 256-timesteps, is summarized in Table 2. The results indicate that both SSIM and PSNR peak at the 16×16 resolution, suggesting that a fourfold super-resolution offers the optimal balance between detail reconstruction and fidelity for the 4×4 input tactile data.

4.4 Texture classification

In order to show the physical significance of super-resolved data from the SuperTac Network, we perform a comparative analysis of texture classification when the original 4×4 tactile data was used for classification compared to super-resolved 16×16 tactile data. We used a CNN based classifier network consisting of two convolution blocks and a softmax layer to classify different types of texture data. We varied the protrusions between bumps and ridges by 3, 4, and 6 to create variation in textured plates. As shown in Figure 11, the classification network yields an accuracy of 76.3% when the 4×4 tactile data was used as input to the classifier, whereas using super-resolved 16×16 tactile data provided a classification accuracy of 93.1%. The 17% jump in texture classification accuracy when super-resolved tactile data was used for classification indicates the ability to capture fine details of textures by the SR tactile data obtained from the SuperTac network.

5 Discussion

The SuperTac framework presents an algorithmic approach to enhancing tactile resolution by leveraging spatiotemporal features of tactile data, enabling the reconstruction of high-resolution outputs from low-resolution inputs. Experimental results demonstrate improvements across multiple metrics, including SSIM, PSNR, and texture classification accuracy while maintaining real-time inference speeds. Unlike conventional approaches that rely on hardware-level enhancements, SuperTac achieves resolution enhancement through software-based processing. Although this introduces a marginal increase in energy consumption due to added computational load, it remains significantly lower than that of hardware-based

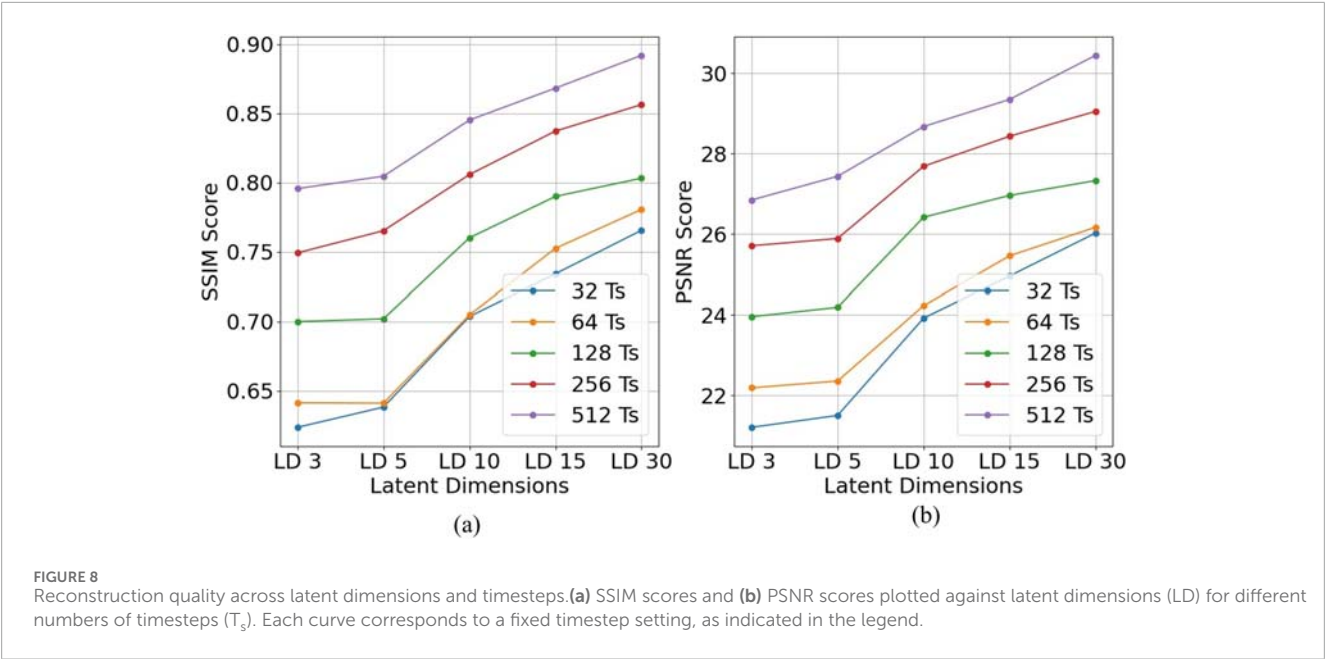
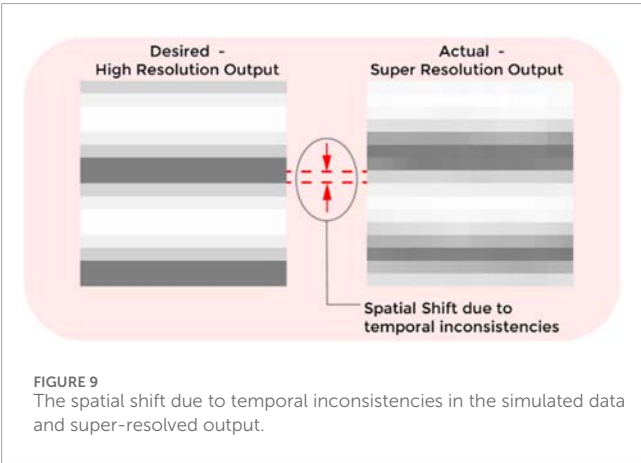


TABLE 1 Keeping the latent dimensions as 10, calculating the metrics SSIM, FPS and Phase Correlation (PC) by varying the timesteps.

Timesteps	SSIM	FPS	PC
512	0.84	48.89	88.76
256	0.80	50.79	94.32
128	0.76	51.24	93.32
64	0.70	51.83	86.27
32	0.70	52.76	85.38



solutions. This software-centric design offers a scalable and cost-effective alternative without modifying sensor hardware, making SuperTac particularly appealing for industrial applications in robotics, prosthetics, and haptic interfaces.

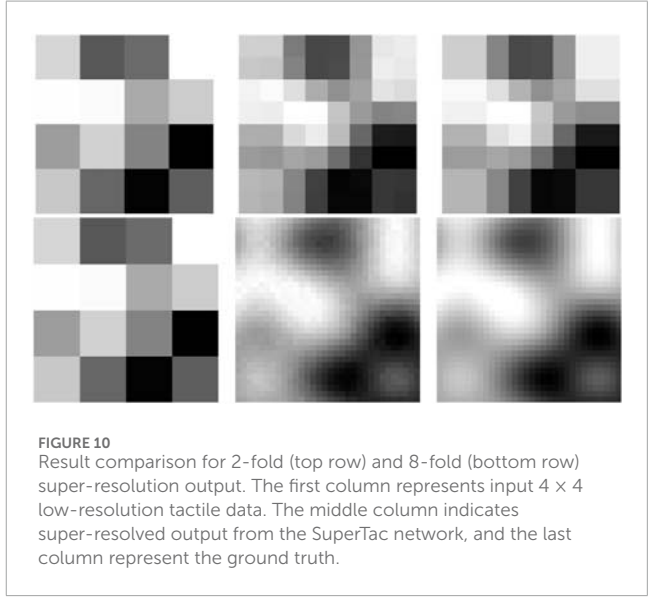


TABLE 2 Average PSNR and SSIM metric comparison for different resolution and constant timestep of 256.

Resolution	PSNR	SSIM
8 x 8	~24.02	~0.77
16 x 16	~27.3	~0.81
32 x 32	~23.09	~0.74

Compared to Wu et al. (2022), who adapted image-based CNNs and GANs for tactile super-resolution (TactileSRCNN and TactileSRGAN), SuperTac achieves comparable resolution

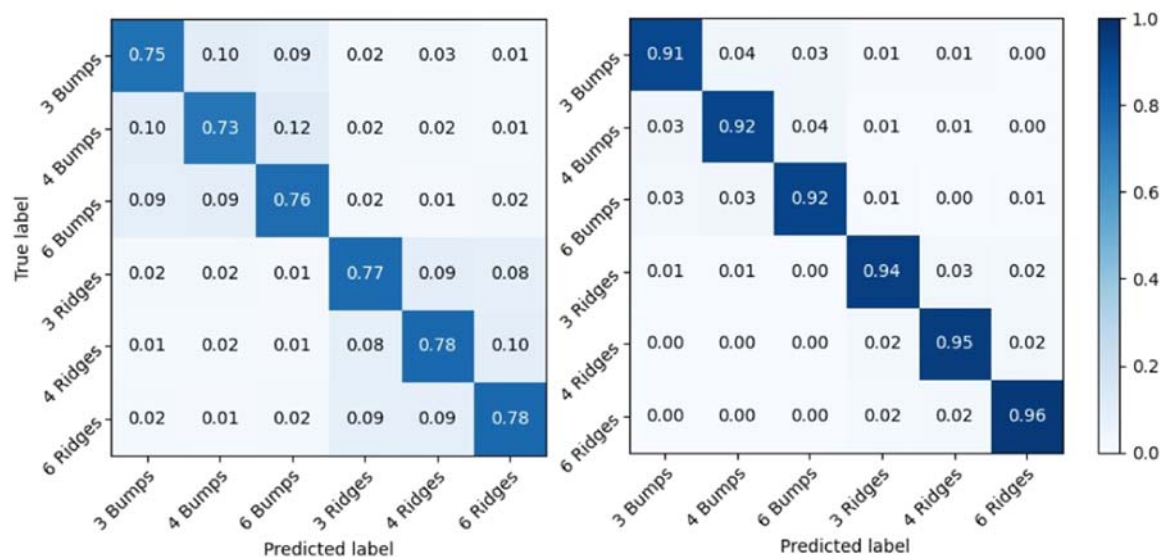


FIGURE 11
Confusion matrix for classification of artificial textures using 4 × 4 (Left panel) and 16 × 16 (Right panel) tactile data using CNN classifier.

enhancement while maintaining real-time processing speeds exceeding 50FPS, which is crucial for online deployment. While Wu et al.'s GAN-based methods provide high perceptual quality, their computational demands are significantly higher, making them less suitable for low-power embedded applications. Similarly, Park et al. (2021) presented EIT-NN, a neural reconstruction framework for EIT-based sensors. Unlike EIT-NN, which is tailored for a specific sensing modality, SuperTac is sensor-agnostic and does not require specialized hardware or calibration procedures, thus offering broader applicability across tactile platforms.

A key factor influencing the performance of the proposed algorithm is the interplay between the latent dimension and the choice of timestep, both of which govern the tradeoff between computational efficiency and output quality. The empirical results show that a latent dimension of 10 provides an effective balance offering near-optimal performance while maintaining high inference speeds suitable for real-time applications. Similarly, the relationship between timesteps and phase correlation (PC) follows a bell-shaped trend rather than a linear one. This can be explained by two competing factors. First, as the timestep decreases (e.g., from 256 to 32), the data resolution decreases, leading to less precise SR image construction, which increases positional error and decreases phase correlation. On the other hand, as the timestep increases (e.g., from 256 to 512), the binned data becomes overly smoothed, losing critical temporal variations needed to accurately capture tactile information, thus also reducing phase correlation. The optimal performance occurs at 256 timesteps, where the balance between data resolution and smoothing leads to the highest phase correlation value (94.325). This corresponds to an average positional error of 0.8 pixels (± 0.65 mm) with respect to the desired high-resolution output. The results in Table 1 illustrate this relationship, emphasizing the importance of selecting an appropriate timestep to optimize the trade-off between resolution and error in tactile data processing.

In this study, we achieved 4 times super-resolution while maintaining the high degree of SSIM and PSNR, there remains potential to further improve performance through a broader and more diverse dataset or by incorporating a denser upsampling network. However, care must be taken to avoid overfitting, particularly as model capacity increases. Therefore, a well-balanced approach to model complexity and dataset diversity, ensuring generalization across various tactile scenarios is necessary.

The performance metrics used in this study are based on empirical evaluations using metrics such as SSIM, PSNR, and texture classification accuracy. To further strengthen the findings, incorporating statistical methods like hypothesis testing (e.g., t-tests or ANOVA) could be beneficial. Although not included in the current study, such methods represent a valuable direction for future work to enhance the statistical rigor and reproducibility of results.

The proposed SuperTac framework is designed to be sensor-agnostic and can, in principle, be applied to different types of tactile sensors. However, certain modifications may be necessary depending on the nature of the sensor data. For example, sensors with higher native resolution or different spatial layouts may require changes in the input encoding or network architecture. Similarly, sensors that capture multi-dimensional data (e.g., force vectors or shear) may benefit from a more complex feature extraction module. Despite these differences, the core idea of using dimensionality reduction to capture spatiotemporal features followed by learned upsampling remains applicable across sensor modalities.

6 Conclusion

In this study, we demonstrated the ability of the SuperTac framework to enhance tactile resolution through a dimensionality reduction network combined with residual upsamplers. The

framework processes spatiotemporal information and generates high-resolution tactile images, as indicated by improved SSIM, PSNR, and FPS metrics. The ability to achieve super-resolution in real-time opens up significant potential for applications in robotics and prosthetics, where high-resolution tactile feedback is crucial for tasks like manipulation and exploration. Despite the promising results, real-world deployment may present additional challenges, including sensor noise, calibration issues, and environmental factors that could affect system performance. These challenges need to be addressed for robust operation in real-world scenarios. The SuperTac framework offers a scalable software-based solution for tactile super-resolution, distinguishing itself from hardware-based approaches used in commercial tactile sensors. This makes it a promising candidate for industry applications in robotics, prosthetics, and other areas requiring enhanced tactile perception. In future work, we plan to explore more diverse datasets and denser upsampling networks to further improve performance. Additionally, real-world validation and statistical analysis could provide deeper insights into the framework's capabilities and potential for broader adoption.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

NP: Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. RR: Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. DK: Conceptualization, Formal Analysis, Investigation, Project administration, Supervision, Writing – original draft, Writing –

review and editing. NT: Conceptualization, Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. To rephrase and refine the language of content originally written by the author(s). All intellectual and scientific content, as well as critical ideas and conclusions presented in this manuscript, are solely the work of the author(s). The authors reviewed and validated all output generated by the AI to ensure accuracy and appropriateness.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abraira, V. E., and Ginty, D. D. (2013). The sensory neurons of touch. *Neuron* 79 (4), 618–639. doi:10.1016/j.neuron.2013.07.051
- Cheng, G., Matsune, A., Li, Q., Zhu, L., Zang, H., and Zhan, S. (2019). "Encoder-decoder residual network for real super-resolution," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*, 2169–2178. doi:10.1109/CVPRW.2019.00270
- Huang, F., Sun, X., Xu, Q., Cheng, W., Shi, Y., and Pan, L. (2025). Recent developments and applications of tactile sensors with biomimetic microstructures. *Biomimetics* 10 (3), 147. doi:10.3390/biomimetics10030147
- Huang, Y., Wang, W., and Wang, L. (2015). Bidirectional recurrent convolutional networks for multi-frame super-resolution. *Adv. Neural Inf. Process. Syst. (NeurIPS)* 28. doi:10.3390/biomimetics10030147
- Jo, Y., Oh, S., Kang, J., and Kim, S. (2018). "Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, Salt Lake City, Utah, USA. New York, NY, USA: IEEE, 3224–3232. doi:10.1109/CVPR.2018.00340
- Johansson, R. S., and Flanagan, J. R. (2009). Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.* 10 (5), 345–359. doi:10.1038/nrn2621
- Kim, M., Choi, H., Cho, K. J., and Jo, S. (2021). Single to multi: data-driven high resolution calibration method for piezoresistive sensor array. *IEEE Robotics Automation Lett.* 6 (3), 4970–4977. doi:10.1109/lra.2021.3070823
- Kingma, D., and Welling, M. (2014). Auto-encoding variational bayes. Available online at: <https://arxiv.org/abs/1312.6114>.
- Kumar, D., Ghosh, R., Nakagawa, A., Soares, A., and Thakor, N. (2020). Neuromorphic approach to tactile edge orientation estimation using spatiotemporal similarity. *Neurocomputing* 407, 181–192. doi:10.1016/j.neucom.2020.04.131
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., et al. (2017). "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, 105–114. doi:10.1109/CVPR.2017.19
- Li, Y., Long, J., Chen, Y., Huang, Y., and Zhao, N. (2022). Crosstalk-free, high-resolution pressure sensor arrays enabled by high-throughput laser manufacturing. *Adv. Mater.* 34 (21), 2200517. doi:10.1002/adma.202200517
- Lu, Z., Su, Y., Ye, Q., Wang, Z., Shi, X., Li, G., et al. (2024). A low-cost super-resolution tactile sensor: design, fabrication and validation. *IEEE Sensors J.* 24 (1), 36518–36529. doi:10.1109/JSEN.2024.3467254
- Meribout, M., Takele, N. A., Derege, O., Rifiki, N., El Khalil, M., Tiwari, V., et al. (2024). Tactile sensors: a review. *Measurement* 238, 115332. doi:10.1016/j.measurement.2024.115332

- Oller, M., Planas i Lisbona, M., Berenson, D., and Fazeli, N. (2023). "Manipulation via membranes: high-resolution and highly deformable tactile sensing and control," in *Conference on robot learning*, 1850–1859. doi:10.48550/arXiv.2209.13432
- Ouyang, Q., Yao, C., Chen, H., Song, L., Zhang, T., Chen, D., et al. (2024). Machine learning-coupled tactile recognition with high spatiotemporal resolution based on cross-striped nanocarbon piezoresistive sensor array. *Biosens. Bioelectron.* 246, 115873. doi:10.1016/j.bios.2023.115873
- Park, H., Park, K., Mo, S., and Kim, J. (2021). Deep neural network based electrical impedance tomographic sensing methodology for large-area robotic tactile sensing. *IEEE Trans. Robotics* 37 (5), 1570–1583. doi:10.1109/tro.2021.3060342
- Parvizi-Fard, A., Amiri, M., Kumar, D., Iskarous, M. M., and Thakor, N. V. (2021). A functional spiking neuronal network for tactile sensing pathway to process edge orientation. *Sci. Rep.* 11 (1), 1320. doi:10.1038/s41598-020-80132-4
- Pyo, S., Lee, J., Bae, K., Sim, S., and Kim, J. (2021). Recent progress in flexible tactile sensors for human-interactive systems: from sensors to advanced applications. *Adv. Mater.* 33 (47), 2005902. doi:10.1002/adma.202005902
- Sankar, S., Brown, A., Balamurugan, D., Nguyen, H., Iskarous, M., Simcox, T., et al. (2019). Texture discrimination using a flexible tactile sensor array on a soft biomimetic finger. *Proc. IEEE Sensors Conf.*, 1–4. doi:10.1109/SENSOR43011.2019.8956704
- Souza, B. C., Lopes-dos Santos, V., Babelo, J., and Tort, A. B. L. (2019). Spike sorting with Gaussian mixture models. *Sci. Rep.* 9 (1), 3627. doi:10.1038/s41598-019-39986-6
- Wang, R., Hu, S., Zhu, W., Huang, Y., Wang, W., Li, Y., et al. (2023). Recent progress in high-resolution tactile sensor array: from sensor fabrication to advanced applications. *Prog. Nat. Sci. Mater. Int.* 33 (1), 55–66. doi:10.1016/j.pnsc.2023.02.005
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., et al. (2018). ESRGAN: enhanced super-resolution generative adversarial networks. Available online at: <https://arxiv.org/abs/1809.00219>.
- Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612. doi:10.1109/TIP.2003.819861
- Wu, B., Liu, Q., and Zhang, Q. (2022). "Tactile pattern super resolution with taxel-based sensors," in *Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 3644–3650. doi:10.1109/IROS47612.2022.9981062
- Yan, Y., Hu, Z., Yang, Z., Yuan, W., Song, C., Pan, J., et al. (2021a). Soft magnetic skin for super-resolution tactile sensing with force self-decoupling. *Sci. Robotics* 6 (51), eabc8801. doi:10.1126/scirobotics.abc8801
- Yan, Z., Wang, L., Xia, Y., Qiu, R., Liu, W., Wu, M., et al. (2021b). Flexible high-resolution triboelectric sensor array based on patterned laser-induced graphene for self-powered real-time tactile sensing. *Adv. Funct. Mater.* 31 (23), 2100709. doi:10.1002/adfm.202100709
- Yu, L., and Liu, D. (2025). Recent progress in tactile sensing and machine learning for texture perception in humanoid robotics. *Interdiscip. Mater.* 4 (2), 235–248. doi:10.1002/idm2.12233
- Zhang, C., Zheng, H., Sun, J., Zhou, Y., Xu, W., Dai, Y., et al. (2022). 3D printed, solid-state conductive ionoelastomer as a generic building block for tactile applications. *Adv. Mater.* 34 (2), 2105996. doi:10.1002/adma.202105996
- Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., and Fu, Y. (2018). Image super-resolution using very deep residual channel attention networks. 294–310. doi:10.1007/978-3-030-01234-2_18