



OPEN ACCESS

EDITED BY

Changxin Gao,
Huazhong University of Science and
Technology, China

REVIEWED BY

Rocco Furferi,
University of Florence, Italy
Longsheng Wei,
China University of Geosciences Wuhan,
China

*CORRESPONDENCE

Guohong Chai
✉ ghchai99@nimte.ac.cn
Guokun Zuo
✉ moonstone@nimte.ac.cn

RECEIVED 15 March 2024

ACCEPTED 20 May 2024

PUBLISHED 30 May 2024

CITATION

Lin Z, Zheng H, Lu Y, Zhang J, Chai G and
Zuo G (2024) Object surface
roughness/texture recognition using machine
vision enables for human-machine haptic
interaction. *Front. Comput. Sci.* 6:1401560.
doi: 10.3389/fcomp.2024.1401560

COPYRIGHT

© 2024 Lin, Zheng, Lu, Zhang, Chai and Zuo.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Object surface roughness/texture recognition using machine vision enables for human-machine haptic interaction

Zixuan Lin^{1,2,3}, Haowei Zheng^{1,2,3}, Yue Lu^{1,2,3}, Jiayi Zhang^{2,3},
Guohong Chai^{2,3*} and Guokun Zuo^{1,2,3*}

¹Cixi Biomedical Research Institute, Wenzhou Medical University, Ningbo, China, ²Laboratory of Advanced Theranostic Materials and Technology, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, Ningbo, China, ³Ningbo Cixi institute of Biomedical Engineering, Ningbo, China

Tactile feedback can effectively improve the controllability of an interactive intelligent robot, and enable users to distinguish the sizes/shapes/compliance of grasped objects. However, it is difficult to recognize object roughness/textures through tactile feedback due to the surface features cannot be acquired with equipped sensors. The purpose of this study is to investigate whether different object roughness/textures can be classified using machine vision and utilized for human-machine haptic interaction. Based on practical application, two classes of specialized datasets, the roughness dataset consisted of different spacing/shapes/height distributions of the surface bulges and the texture dataset included eight types of representative surface textures, were separately established to train the respective classification models. Four kinds of typical deep learning models (YOLOv5l, SSD300, ResNet18, ResNet34) were employed to verify the identification accuracies of surface features corresponding to different roughness/textures. The human fingers' ability to objects roughness recognition also was quantified through a psychophysical experiment with 3D-printed test objects, as a reference benchmark. The computation results showed that the average roughness recognition accuracies based on SSD300, ResNet18, ResNet34 were higher than 95%, which were superior to those of the human fingers (94% and 91% for 2 and 3 levels of object roughness, respectively). The texture recognition accuracies with all models were higher than 84%. Outcomes indicate that object roughness/textures can be effectively classified using machine vision and exploited for human-machine haptic interaction, providing the feasibility of functional sensory restoration of intelligent robots equipped with visual capture and tactile stimulation devices.

KEYWORDS

human-machine interaction, roughness/texture recognition, machine vision, deep learning model, tactile feedback

1 Introduction

Currently, there is still a wide gap between the functionality and utility of the robots and the users' expectations. An acceptable level of intention control function can be obtained by translating the muscle electrical activities of peripheral limbs (Farina et al., 2023) or the central nerve activities of the brain (Luo et al., 2023) into the commands

of an intelligent robotic device via multi-channel neural control interfaces (Makin et al., 2023). However, the existing interactive robots generally have common shortcomings, such as poor performance in interactive control, insufficient ability in human-machine collaboration, and a lack of active sense capability (Yang et al., 2019). Lack of tactile feedback has been regarded as one of the major drawbacks (Jabban et al., 2022). To date, except for a few examples with limited application (Svensson et al., 2017; Sensinger and Dosen, 2020), no commercially available robot can provide effective tactile feedback for its users to close the interactive control loop (Jabban et al., 2022).

Tactile sensory feedback undertakes an unparalleled role when humans interact with external surroundings due to its realistic sense of presence (Sigrist et al., 2013; Svensson et al., 2017; Sensinger and Dosen, 2020). Numerous studies have confirmed that providing tactile feedback can not only improve the controllability (George et al., 2019; Sensinger and Dosen, 2020) and the sense of embodiment of robotic devices (Rosén et al., 2009; Page et al., 2018), but also promote human-machine motor coordination (Clemente et al., 2019; Weber and Matsiko, 2023) and manual efficiency (George et al., 2019; Chai et al., 2022). Tactile feedback can be elicited on a human subject by selectively stimulating his/her peripheral nervous system using non-invasive or invasive stimulation approaches (Sensinger and Dosen, 2020; Jabban et al., 2022). The prior closed-loop feedback studies displayed that the sensing information of an intelligent robot (e.g., grip force or aperture angle of a prosthetic hand, etc.) could be delivered to its users by modulating stimulation parameters or/and changing the active stimulation units (e.g., a set of electrode pads or vibrators) (George et al., 2019; Chai et al., 2022). The users were able to regulate the grip force or the grip (line/angular) displacement of a robot device to perform delicate object manipulation or recognize the size/compliance of the grasped objects through the delivered tactile sensations (George et al., 2019). Hence, it is essential to develop effective tactile feedback from an intelligent robot to its users to improve the coordination in the manipulation of robotic devices and increase the presence of human-machine interaction.

It has been proven that tactile feedback enables fine grip force control of robotic devices (George et al., 2019; Chai et al., 2022) and users' attention allocation (Sklar and Sarter, 1999; Makin et al., 2023) and confers to the users the ability to recognize object size and compliance, with or without visual feedback (Arakeri et al., 2018; George et al., 2019; Chai et al., 2022). However, in terms of human-machine interaction, so far, the 2 types of typical object physical properties, surface roughness and texture, were unable to be delivered to the users through a haptic interface due to the complexity of the surface features (Bajcsy, 1973; Hashmi et al., 2023). A key reason is that the current contact sensors (e.g., pressure/ displacement transducers) cannot identify and differentiate a wide variety of features of the surface roughness and texture. The previous related studies have proved that providing tactile feedback for object roughness/texture recognition has important application requirements in a variety of fields ranging from medical rehabilitation to education and entertainment (Samra et al., 2011; Lin and Smith, 2018), especially roughness/texture recognition for blind/visually impaired patients (O'Sullivan et al., 2014).

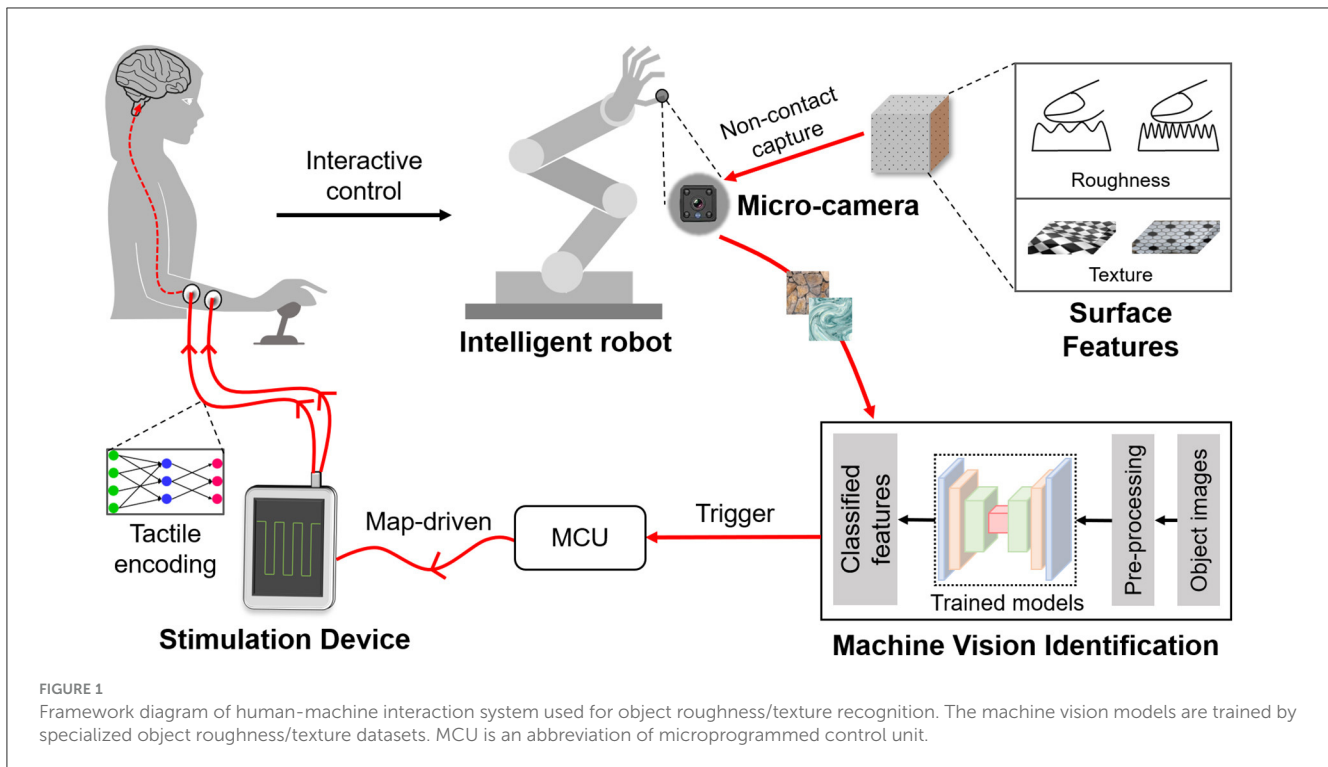
Machine vision has emerged as the most promising non-contact technique for object detection and image classification by combining cameras, videos, and deep-learning methods (Zhao et al., 2019; Georgiou et al., 2020; Hashmi et al., 2023). Recent representative studies have shown that compared to conventional methods, machine vision-based approaches could better assess the surface quality of machined parts by contactlessly measuring the surface characteristics, i.e., surface roughness, waviness, flatness, surface texture, etc. (Hashmi et al., 2023). The image calibration algorithm based on Bayes theorem has been proved can effectively restore the blurred images of object surface roughness (Dhanasekar and Ramamoorthy, 2010). The image-based machine vision model not only can solve the measurement of surface characteristics, but also avoid the shortcomings (e.g., limited assessment area/position, limited sensor precision, etc.) of traditional contact measurement methods. Moreover, in contrast to the manual and traditional machine learning methods, the deep learning model are also widely used for pavement texture recognition (Chen et al., 2022) and fingerprint classification (Mukoya et al., 2023) due to its powerful training capabilities and portability. In addition, electrotactile stimulation is extensively applied for sensory restoration of typical intelligent robotic devices (e.g., prosthetic hands.) due to the apparent advantages, such as, non-invasive, multi-parameter adjustable, and compact electronics (Svensson et al., 2017; Sensinger and Dosen, 2020). Therefore, machine vision combined with electrotactile feedback may be a viable alternative for object roughness/texture recognition. Whether different object roughness and textures could be effectively discriminated with deep learning models deserves to be investigated in depth.

The purpose of this study is to investigate the capability of machine vision to recognize object roughness and texture. We hypothesized that machine vision identifying object roughness/texture could be used for human-machine haptic interaction. To this aim, 2 classes of specialized roughness and texture datasets were separately established to train and test the deep learning models. We employed four types of typical deep learning algorithm models (YOLOv5l, SSD300, ResNet18, and ResNet34) to evaluate the identification performance of machine vision. The human roughness recognition ability was also quantified as the basal control. The outcomes provide important insights into object roughness/texture recognition through tactile feedback. Machine vision combined with electrotactile feedback has the potential to be directly applied to human-machine haptic interaction.

2 Materials and methods

2.1 The human-machine interaction framework for object roughness/texture recognition

The framework of the human-machine interaction system by combination of machine vision and tactile feedback, used for object roughness/texture recognition, is illustrated in Figure 1. It consists of interactive control and image capture, object roughness/texture classification, and tactile recognition, in terms of the interactive function. The users could manipulate an



intelligent robot to capture the surface roughness/texture images of targeted objects via an integrated non-contacting micro-camera. Different object roughness/textures could be distinguished by classifying the roughness/texture features of the acquired images through a specialized machine vision identification model. The classification results, transmitted in the forms of different category labels, were adopted to trigger a (mechanical/electrical) stimulation device to output different stimulus pulses to evoke distinguishable tactile sensations on users (through surface stimulation electrodes/vibrators), via an MCU with pattern-mapped driven protocols. The users were able to recognize object roughness/texture through tactile feedback. Since it has been proven that users can effectively identify tactile feedback through different encoding strategies (Achanccaray et al., 2021; Tong et al., 2023), how to classify object roughness/texture features and train the machine vision identification model are critical for human-machine haptic interaction.

2.2 Deep learning models

In the current study, we adopted four typical deep learning models, YOLOv5l (Glenn-jocher and Sergiuwaxmann, 2020), SSD300 (Liu et al., 2016), ResNet18, and ResNet34 (He et al., 2016), which have been widely used in the field of machine vision (Arcos-García et al., 2018; Chen et al., 2021; Jiang et al., 2022), to classify the features of object roughness and textures. The feature classification capabilities of the models were trained and tested by two constructed object roughness and texture datasets. The initialization parameters of these models are presented in Table 1.

TABLE 1 The initialization parameters of 4 types of deep learning models.

Models	Parameters	Value
YOLOv5l	image size	640 × 640
	Batch size	16
	Learning rate	0.01
	Epochs	200
SSD300	image size	300 × 300
	Batch size	4
	Learning rate	0.0001
	Iterations	400000
	Weight decay	0.005
	Momentum	0.9
ResNet18/34	image size	640 × 363
	Batch size	16
	Learning	0.1
	Epochs	700

2.3 Object roughness/texture datasets

In order to verify the utility of the selected machine learning models, object surface roughness and texture datasets were established, separately. We calculated the mean and standard deviation of 2 datasets to achieve image equalization and normalization. Then we used the hold-out method to randomly

divide 2 datasets into the training set and test set with a ratio of 9:1 to test the models, respectively.

2.3.1 Surface roughness (SR) dataset

The images of the roughness dataset are acquired from the custom-made objects for this study, as shown in [Figure 2](#). Each test object was composed of a square basal block (50 mm × 50 mm × 10 mm) and the part of the roughness bulges (1 mm in the circumference). According to the requirements of roughness identification, we established the image dataset by changing the combinations of 3 characteristics of the surface bulges, i.e., the heights, spacing, and cross-sectional shapes of the bulges. Taking the performance of haptic interface and human tactile recognition ability into consideration, according to the prior study, the human tactile recognition ability of roughness peaks at a spacing of 3.2 mm or 4.5 mm between bulges and then declines for increases in bulge spacing up to 6.2 mm or 8.8 mm, respectively, in different directions of the scan ([Connor et al., 1990](#)). Thus, 5 levels of spacing (i.e., SP1 = 3 mm, SP2 = 4.5 mm, SP3 = 6 mm, SP4 = 7.5 mm, SP5 = 9 mm), 5 kinds of cross-sectional shapes (circle, triangle, square, pentagon, hexagon), and 2 types of height arrangement modes (rows and rings-shaped of high-low heights) were adopted to construct different surface roughness. The bulge heights were, respectively, 10 mm and 5 mm for the high and low patterns, the aim being to avoid contact between the digit and the basal block between adjacent bulges.

The entire SR dataset encompassed 50 roughness categories according to the selected combinations of three bulge characteristics. Each category was expanded to 30 images with different perspectives through the coordinate transformation of image rotation, scaling, flipping, and affine transformation. There were 1,500 roughness images of the surface roughness in the entire SR.

2.3.2 Surface texture (ST) dataset

The ST dataset was acquired from the Describable Textures Dataset (DTD) ([Cimpoi et al., 2014](#)). The DTD encompassed 47 texture categories, characterized by vocabulary-based texture attributes. All images were collected from Google and Flickr's websites by searching for relevant attributes and terms.

Similarly, taking into account the recognition capability of the tactile feedback interface and the utility of training models, we selected images with eight categories of texture features (dotted, striped, checkered, cracked, pure, swirly, honeycombed, and scaly) from the DTD. Each category contained 120 images, and there were 960 images in the full ST dataset. This ST dataset selected texture features that are representative in terms of haptic interactions, as visually illustrated in [Figure 3](#).

2.4 Human roughness recognition experiment

This experiment aims to quantify human's ability to discriminate object roughness through tactile feedback, serving as

a reference benchmark for evaluating the identification capabilities of the current deep learning models.

2.4.1 Subjects

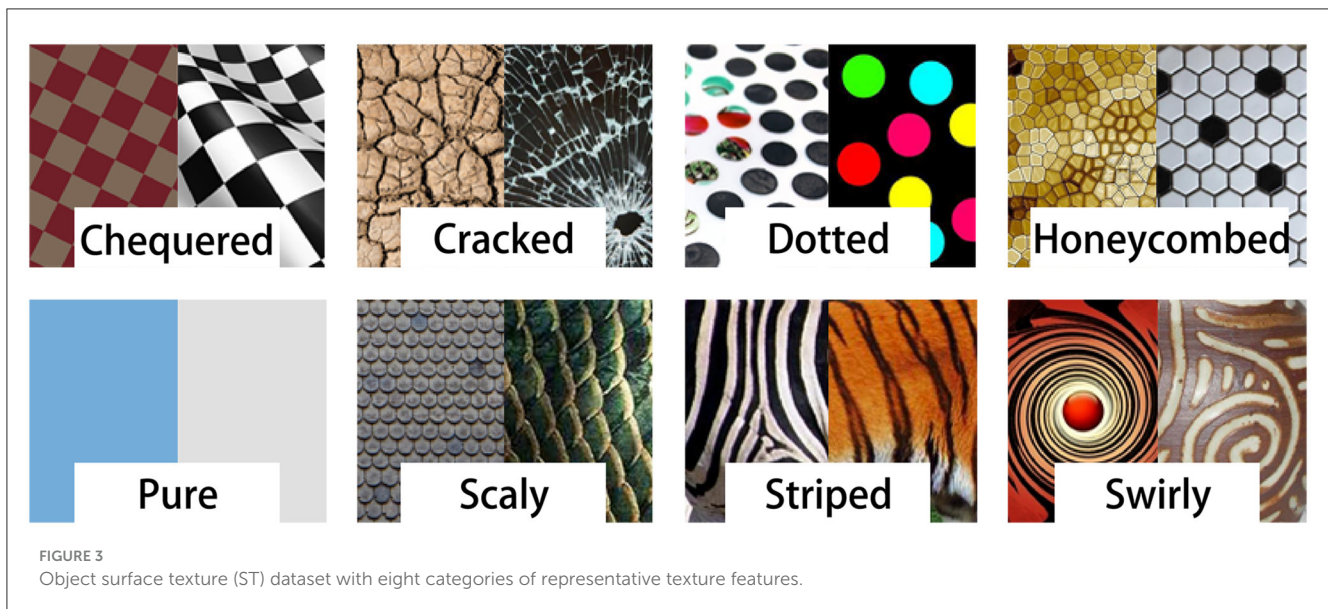
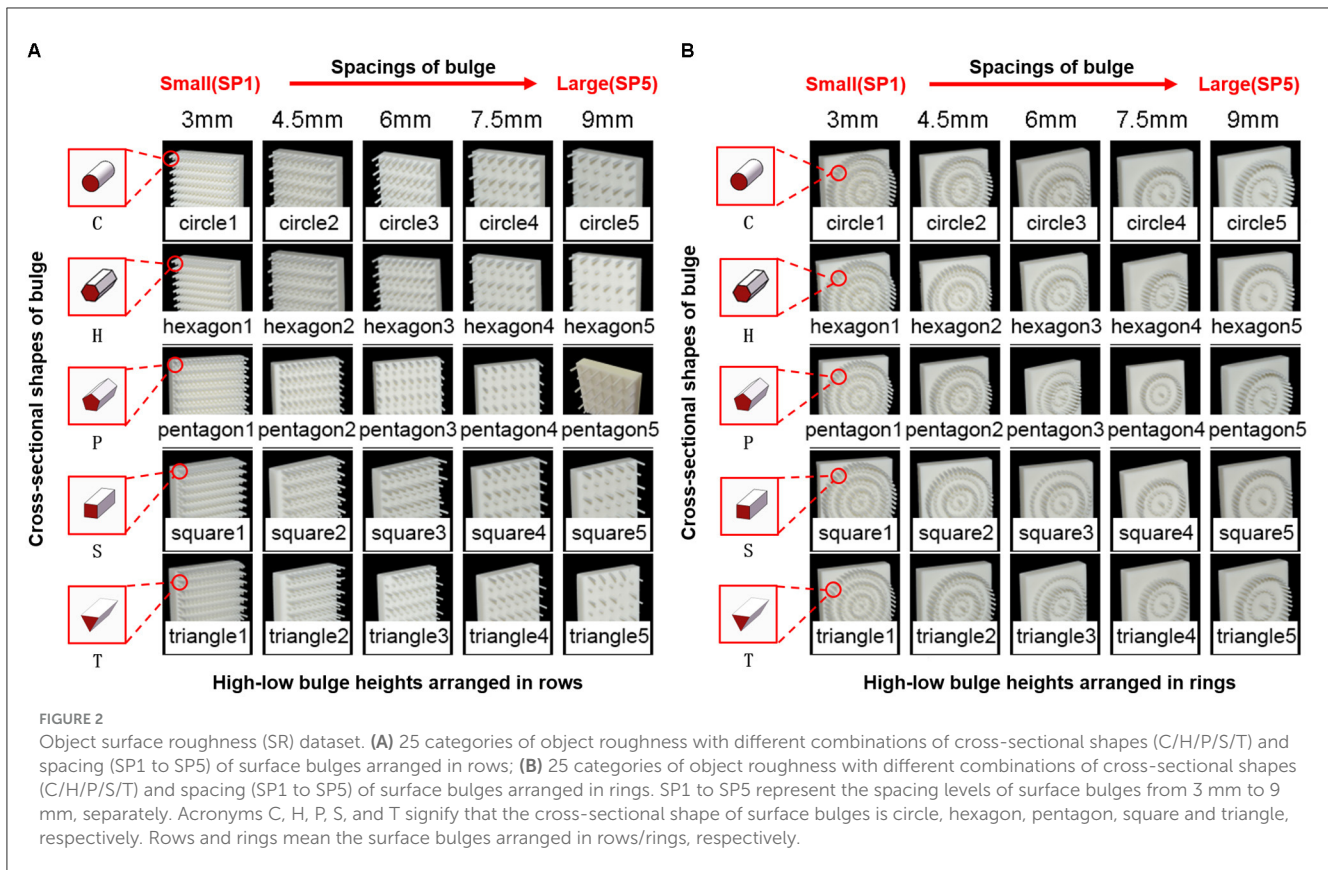
9 healthy, able-bodied subjects (age 22-25 years, 5 females, all right-handed) participated in the study. All experiments were conducted in accordance with the latest version of the Declaration of Helsinki and approved by the Ethics Committee of Human and Animal Experiments of Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences. All subjects were informed about the experimental procedure and signed the informed consent forms prior to participation.

2.4.2 Experimental procedure

The 3D-printed objects with preset roughness ([Figure 2](#)) were chosen as the test objects. Since the same cross-sectional shapes of the surface bulges cannot be distinguished by the finger's touch (prior test), we selected these objects with 10 categories of object roughness (five spacing × two height arrangements × circular cross-sectional of surface bulges). The whole experiment was divided into four different experimental groups according to different recognition tasks and test objects, as shown in [Table 2](#). In each group, 2 or 3 objects were randomly selected from 10 test objects as the identification targets. Every subject, whose visual and auditory cues were isolated with a sleep mask and a pair of noise-canceling headphones playing gray noise, was instructed to distinguish the 2 or 3 kinds of surface roughness (objects) with his/her right index finger, respectively. To ensure objective quantification, every test object was provided three times in each group in a pseudo-random order. The in-group sequences of the test objects, inter-group order and the test timing were conducted based on E-Prime 2.0 software on a laptop computer. In each trial, in accordance with the pre-defined sequence, 2 or 3 target objects were well placed in turn at a time, the experimenter signaled the subject to begin touching (identifying) by tapping his/her shoulder, and synchronously pressed the space key on a keyboard to record the starting time. Once 2 or 3 kinds of object roughness could be identified, the subject was asked to press number keys 2 or 3 on a keyboard with his/her contralateral hand to stop tactile recognition, and made the laptop record the recognition time. Then the subject verbally reported the answers (roughness category and the specific roughness levels (if 3 target objects were distinguished)) to the experimenter. The experimenter just input the subject's answer to the laptop and began the next trial. The duration of the overall experiment was about 1 hour. A break of 1-5 minutes was randomly given between different test trials or inter-groups to allow the subjects to relax and keep good physical condition.

2.5 Statistical analysis

Statistical analysis of the data was performed using SPSS 23.0. A paired sample T-test was employed to measure the significant differences in recognition performances among different models or across distinct roughness/texture categories, respectively. A significance level of $p < 0.05$ was considered statistically significant.



3 Results

3.1 Identification performance of machine vision

3.1.1 Object roughness identification

4 types of deep-learning models were trained and tested by the proposed SR dataset. The specific identification accuracies for

single or multiple features of surface roughness are illustrated in Figure 4 and summarized in Table 3. As shown in Table 3, the average accuracies of the YOLOv5l, SSD300, ResNet18, and ResNet34 models were $84.4\% \pm 3.1\%$, $98.6\% \pm 0.9\%$, $95.3\% \pm 1.9\%$, and $96.7\% \pm 1.4\%$, respectively. The paired T-test demonstrated that SSD300 and ResNet18/34 exhibited significantly superior accuracies in roughness identification when compared to YOLOv5l ($p < 0.01$).

TABLE 2 Test objects used for human roughness discrimination.

Group	Number of select/total objects	Characteristics of surface bulges		
		Height arrangement	Spacing (mm)	Shape
a	2/10	In rows	3/4.5/6/7.5/9	Circle
b	3/10	In rows	3/4.5/6/7.5/9	Circle
c	2/10	In rings	3/4.5/6/7.5/9	Circle
d	3/10	In rings	3/4.5/6/7.5/9	Circle

Height arrangement: the surface bulges with two (high/low) distinct heights were alternatively arranged in rows or rings; Spacing: the spacing levels of surface bulges; Shape: the cross-sectional shape of surface bulges.

The identification capacities for 3 categories of surface bulges are depicted in Figure 4. Overall, the four types of models exhibited relatively higher accuracies for bulge spacing and height arrangement discrimination than those for cross-sectional shape. Specifically, the 4 types of models were able to identify the bulge spacing with an average accuracy of $96.1\% \pm 1.5\%$, $100\% \pm 0\%$, $99.3\% \pm 0.6\%$, and $99.3\% \pm 0.6\%$; to identify the bulge cross-sectional shape with an average accuracy of $87.0\% \pm 5.1\%$, $98.6\% \pm 0.9\%$, $98.0\% \pm 2.0\%$, and $98.0\% \pm 1.4\%$; to identify the bulge height arrangement with an average accuracy of $98.7\% \pm 1.5\%$, $100\% \pm 0\%$, $98.0\% \pm 2.0\%$, and $99.3\% \pm 0.5\%$, respectively. The statistical results revealed that no significant differences were found among the 4 types of models in identifying the 3 categories of bulge characteristics ($p \geq 0.05$), except for YOLOv5l for cross-sectional shape ($p < 0.01$).

To compare the human fingers' roughness recognition ability, Table 4 presents the average identification accuracies of the 4 types of models for these objects with circular cross-sectional bulges. For YOLOv5l, SSD300, ResNet18, and ResNet34, the overall identification accuracies of the 4 models were 89.7%, 100%, 100%, and 96.7%, respectively. Specifically, the accuracies for bulge spacing were 93.1%, 100%, 100%, and 96.7%, separately, while the accuracies for height arrangement of bulges were 100% for all models.

3.1.2 Object texture identification

The recognition outcomes of the 4 models for 8 types of textures are presented in Figure 5, Table 5. As illustrated in Figure 5, Four kinds of deep learning models led to similar identification performance across the eight types of textures. Specifically, the average identification accuracies of surface textures were 88.8% for YOLOv5l, 86.0% for SSD300, 84.2% for ResNet18, and 86.3% for ResNet34, respectively. No significant difference appeared between recognition accuracies among four types of models.

3.2 Tactile recognition performance of object roughness

The tactile recognition accuracies for surface roughness with selected objects (Table 2) are displayed in Figure 6, Table 6. The confusion matrices (Figure 6) displayed that the recognition errors largely happened at adjacent roughness levels, irrespective of 2-object (i.e., SP3/SP4 and SP4/SP5) or 3-object (i.e.,

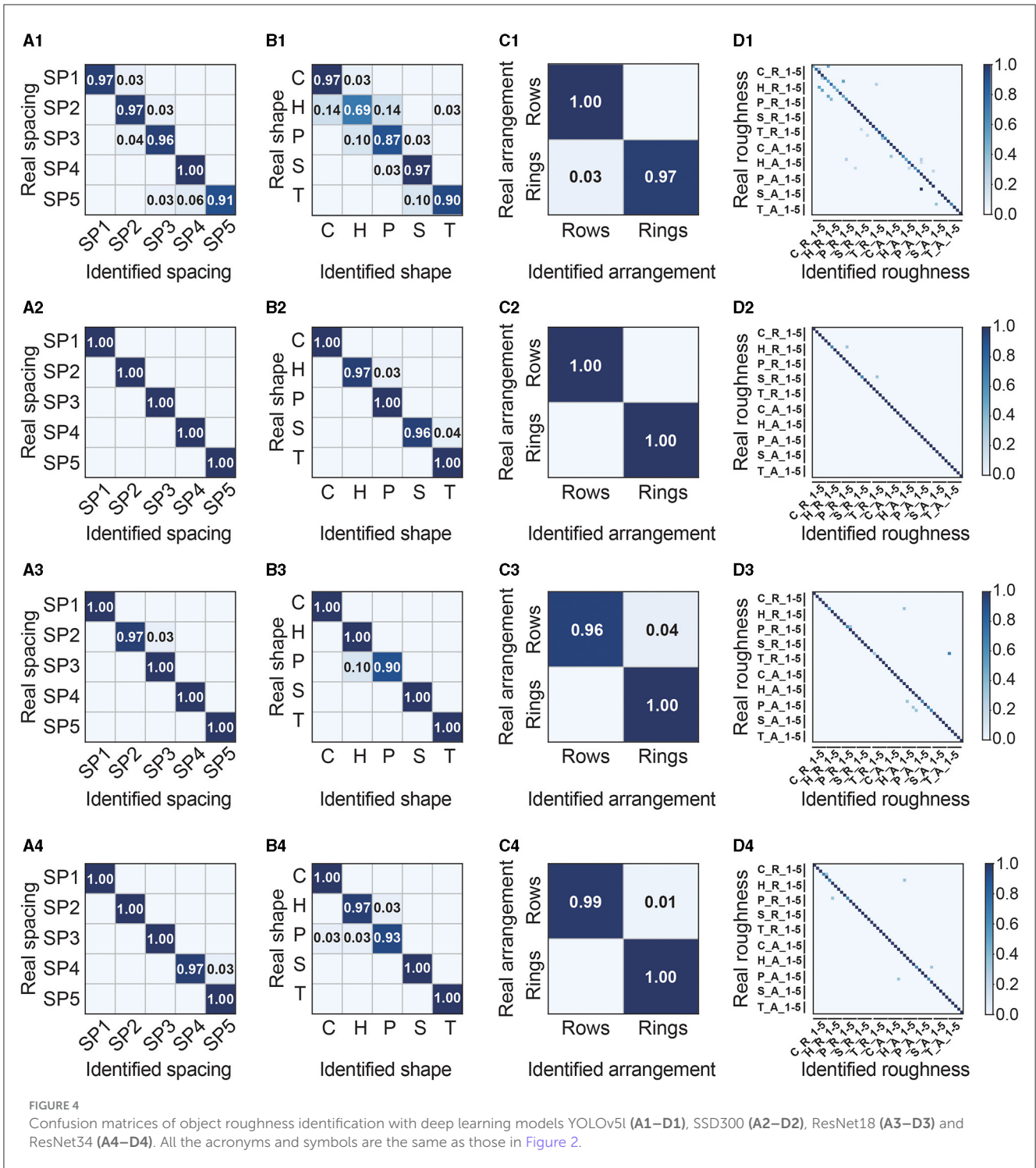
SP2/SP3/SP4 and SP3/SP4/SP5) recognition. The recognition accuracies (Table 6) for 2 levels of roughness were significantly higher than those of 3 levels of roughness ($p < 0.05$). No significant difference appeared between types of bulge height arrangements.

The tactile recognition time for object roughness with selected objects (Table 2) was summarized in Table 7. For groups a,b,c, and d, the average recognition time was $1.5s \pm 0.8s$, $2.0s \pm 1.1s$, $1.7s \pm 0.9s$, and $2.1s \pm 1.1s$, respectively. The paired-sample T test showed that there were significant differences between 2-level and 3-level roughness recognition, and two types of bulge height arrangement ($p < 0.05$), respectively. Notably, significant differences in tactile recognition time were observed when recognizing between 2 or 3 objects and bulges arranged in rows or rings ($p < 0.05$). In particular, the roughness recognition time with adjacent bulge spacing was significantly longer than those with non-adjacent spacing ($p < 0.05$).

4 Discussion

This study sought to demonstrate that object roughness/texture identified via machine vision could be utilized for human-machine haptic interaction. The results of calculating and analyzing revealed that 4 types of models trained with specialized roughness/texture datasets could classify the features of object roughness and texture into the expected categories at favorable accuracies. These outcomes suggest the feasibility of machine vision recognizing object roughness/texture being used for human-machine tactile recognition.

Prior studies have shown that machine vision could provide reliable image-based solutions to inspect object roughness by distinguishing the characteristics of the surface bulges (Dhanasekar and Ramamoorthy, 2010; Jeyapooan and Murugan, 2013; Hashmi et al., 2023). In terms of bidirectional human-machine haptic interaction, how to categorize the object roughness and translate the classification results into effective tactile feedback (sensations) on the users is the key to object roughness recognition (Connor et al., 1990; Lin and Smith, 2018). Therefore, we established a specialized SR dataset to train and test four types of typical models. As shown in Table 3, the test results showed that 4 kinds of deep learning models could identify the bulge characteristics with considerable accuracies ($>84.4\%$). Specifically, models SSD300, ResNet18/34 had better identification performances than those of YOLOv5l both in single and combined characteristics. It can be interpreted that SSD300 employs a multi-scale candidate box



detection algorithm that can effectively improve the generalization ability of feature recognition with different aspect ratios, achieving a high recognition rate for intensive features (Liu et al., 2016). ResNet adopts shortcut connections, effectively preventing the degradation problem caused by network layers increased (He et al., 2016), guaranteeing the classification precision. The difference between ResNet18 and ResNet34 is the number of network layers. By comparison, despite with relatively lower identification accuracies, YOLOv5l might also be applied for roughness

recognition due to its fast detection speed, and it is good at object detection on videos and camera-captured images. The calculation results denote that different object roughness, represented by different characteristic combinations of surface bulges, can be effectively identified by machine vision and applied to human-machine haptic interaction. The SSD300 may be preferred for roughness recognition. Nevertheless, in actual functional tasks, different algorithm models can be flexibly selected according to different task difficulties and requirements.

Furthermore, we quantified the human ability to object roughness recognition using four groups of test objects with specialized roughness characteristics (Table 2), as a benchmark reference to the identification capability of machine vision. The quantified results illustrated that the subjects could correctly discriminate 2 and 3 categories of object roughness at favorable accuracies (Table 6) and recognition time (Table 7). The variation

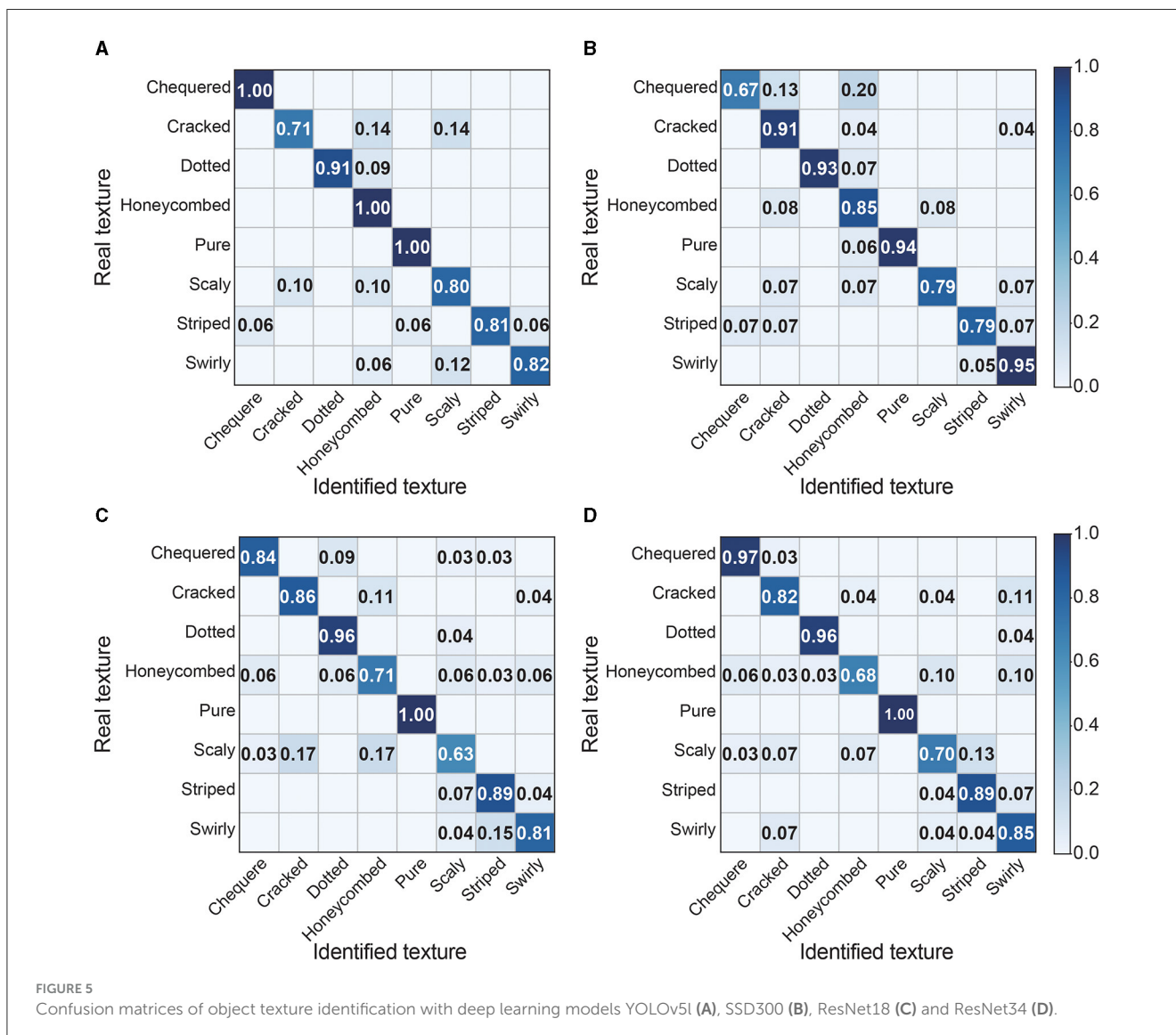
ranges of average recognition rate were less than 4% across 9 randomly recruited subjects. It indicates that all subjects can effectively distinguish the current object surface roughness without significant differences, and it was not essential to add additional subjects in current study. The recognition rates of 2 categories of roughness with different bulge spacings were slightly higher than those of the three categories, and the misidentification

TABLE 3 Summary of overall identification results of SR dataset with 4 kinds of deep learning models.

Models	Spacing (%)	Shape (%)	Arrangement (%)	Avg. (%)
YOLOv5l	96.1±1.5	87.0±5.1	98.7±1.5	84.4±3.1
SSD300	100.0±0.0	98.6±0.9	100.0±0.0	98.6±0.9
ResNet18	99.3±0.6	98.0±2.0	98.0±2.0	95.3±1.9
ResNet34	99.3±0.6	98.0±1.4	99.3±0.5	96.7±1.4

TABLE 4 Summary of overall identification results of object roughness (circular cross-sectional shape) with 4 kinds of deep learning models.

Models	Shape	Spacing (%)	Arrangement (%)	Avg. (%)
YOLOv5l	Circle	93.1	100.0	89.7
SSD300	Circle	100.0	100.0	100.0
ResNet18	Circle	100.0	100.0	100.0
ResNet34	Circle	96.7	100.0	96.7



mainly appeared between the objects with adjacent bulge spacings (Figure 6). This can be explained by the fact that the human perception of surface roughness is the function of a spatial variation code mediated by the activation of mechanoreceptor afferents, especially the SA-I afferents (Connor et al., 1990; Johnson and Hsiao, 1994; Chapman et al., 2002). The human psychophysical recognition results are highly correlated with the population discharge of SA-I afferents (Connor et al., 1990). A smaller bulge spacing on a certain object surface could activate a higher discharge rate and the adjacent discharge rates are likely to result in the decline of the recognition accuracies. Meanwhile, although with different recognition time, the subjects could recognize 2 kinds of bulge height arrangements without significant differences (Tables 6 and 7). It may due to the two kinds of height arrangements is not challenging enough to detect the roughness differences. More classes of height arrangements are still necessary to further evaluate human roughness recognition ability. Accordingly, we calculated the overall identification accuracies of bulge characteristics of the identical test objects using the four kinds of models (Table 4).

TABLE 5 Summary of overall identification results of ST dataset with 4 kinds of deep learning models.

Models	Average accuracy(%)
YOLOv5l	88.8±4.0
SSD300	86.0±3.5
ResNet18	84.2±4.3
ResNet34	86.3±4.3

The summary results displayed that the machine vision was not only able to distinguish different bulge spacings and height arrangements with relatively higher accuracies (Table 4), but also could identify the different cross-sectional shapes of the bulges which were hard to distinguish by touch perception at excellent accuracies (Table 3 and Figure 4). Therefore, it can be concluded that identifying the bulge characteristics with suitable deep learning models appears to be a viable approach for object roughness recognition, which may be sufficient for a wide range of functional human-machine tasks.

Humans cannot intuitively perceive the surface texture of an object if the visual feedback is unavailable. How to transfer different surface textures to the users through tactile feedback has important academic and application value, such as the TACTICS (Way and Barner, 1997), the automatic assigning of haptic texture models (Hassan et al., 2017), and the haptic texture rendering method based on the adaptive fractional difference (Hu and Song, 2022). The feasibility of recognizing object surface features via machine vision for human-machine haptic interaction was also evidenced by the results of object texture classification and identification (Figure 5, Table 5). The calculation results exhibited that four types of machine learning models could identify eight categories of representative texture features at acceptable accuracies (>84%). This indicates that machine vision is able to be effectively exploited to distinguish object texture. The current average accuracy was about 15% higher than the reported results of the DTD (Cimpoi et al., 2014). This is mainly because we only preset 8 categories of object texture labels in the ST dataset in consideration of the limitations of the current haptic interface performance (Svensson et al., 2017; George et al., 2019; Sensinger and Dosen, 2020;

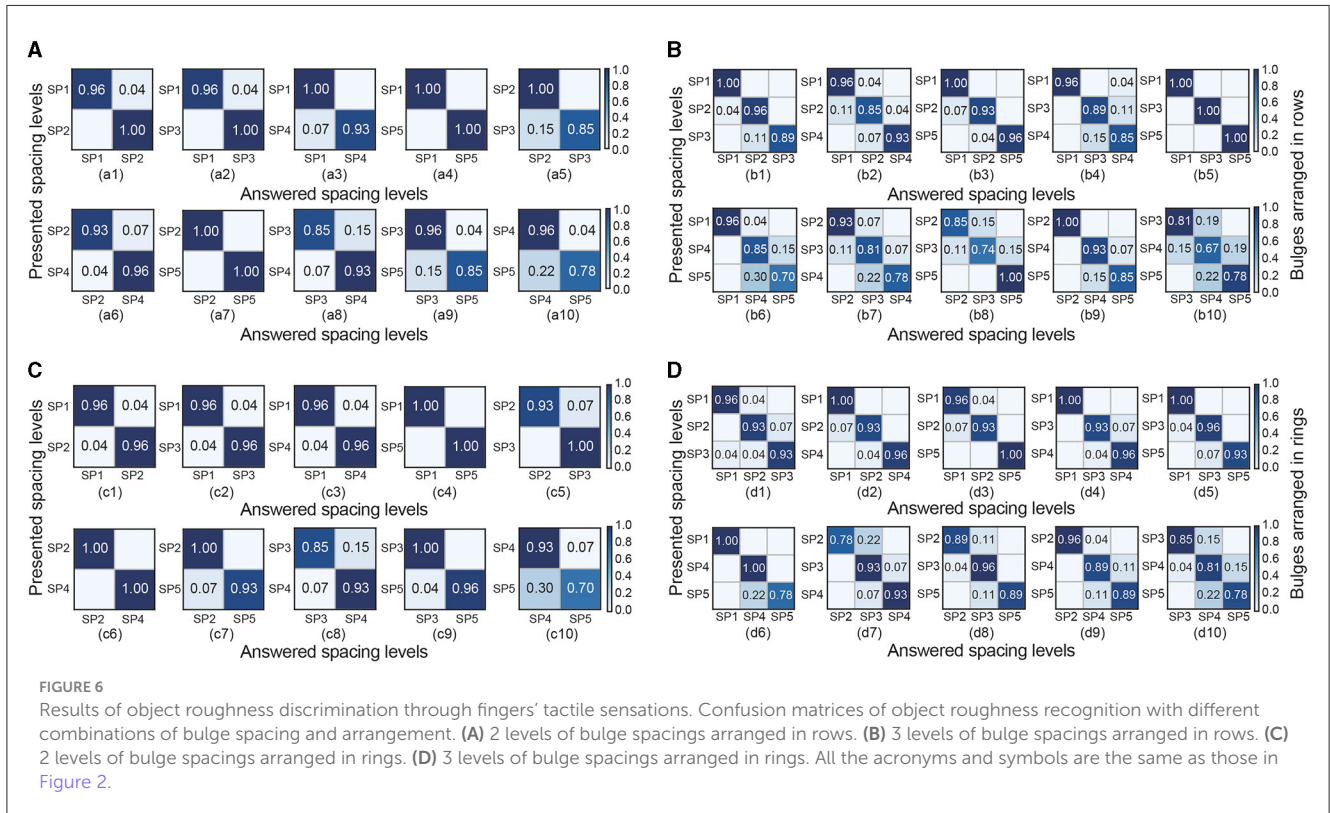


FIGURE 6 Results of object roughness discrimination through fingers' tactile sensations. Confusion matrices of object roughness recognition with different combinations of bulge spacing and arrangement. (A) 2 levels of bulge spacings arranged in rows. (B) 3 levels of bulge spacings arranged in rows. (C) 2 levels of bulge spacings arranged in rings. (D) 3 levels of bulge spacings arranged in rings. All the acronyms and symbols are the same as those in Figure 2.

TABLE 6 Summary of recognition accuracies of object roughness through fingers' tactile perception.

Group*	SP1/2(%)	SP2/3(%)	SP3/4(%)	SP4/5(%)	SP1/3(%)	SP2/4(%)	SP3/5(%)	SP1/4(%)	SP2/5(%)	SP1/5(%)	Avg.(%)
a	98.1±2.0	92.6±7.5	88.9±4.0	87.0±9.0	98.1±2.0	94.4±1.5	90.7±5.5	96.3±3.5	100.0±0.0	100.0±0.0	94.6±3.5
c	96.3±0.0	96.3±3.5	88.9±4.0	81.5±1.2	96.3±0.0	100.0±0.0	98.1±2.0	96.3±0.0	96.3±3.5	100.0±0.0	95.0±2.5
Group*	SP1/2/3(%)	SP2/3/4(%)	SP3/4/5(%)	SP1/2/4(%)	SP2/3/5(%)	SP1/4/5(%)	SP2/4/5(%)	SP1/2/5(%)	SP1/3/5(%)	SP1/3/4(%)	Avg.(%)
b	95.1±3.2	84.0±4.6	75.3±4.3	91.4±3.3	86.4±7.5	84.0±7.5	92.6±4.3	96.3±2.0	100.0±0.0	90.1±3.2	89.5±4.0
d	93.8±1.0	87.7±5.0	81.5±2.0	96.3±2.0	91.4±2.3	92.6±7.3	91.4±2.3	96.3±2.0	96.3±2.0	96.3±2.0	92.3±2.8

*Groups a, b, c and d are the same as those in Table 2.

TABLE 7 Summary of recognition time of object roughness through fingers' tactile perception.

Group*	SP1/2(s)	SP2/3(s)	SP3/4(s)	SP4/5(s)	SP1/3(s)	SP2/4(s)	SP3/5(s)	SP1/4(s)	SP2/5(s)	SP1/5(s)	Avg.(s)
a	1.2±0.4	1.7±0.8	2.0±0.9	2.2±1.0	1.1±0.4	1.6±0.6	1.6±0.7	1.0±0.5	1.1±0.3	1.1±0.4	1.5±0.8
c	1.8±0.8	2.0±0.7	2.0±1.3	2.0±0.9	1.5±0.7	1.4±0.4	2.0±0.9	1.6±1.1	1.5±0.6	1.3±0.7	1.7±0.9
Group*	SP1/2/3(s)	SP2/3/4(s)	SP3/4/5(s)	SP1/2/4(s)	SP2/3/5(s)	SP1/4/5(s)	SP2/4/5(s)	SP1/2/5(s)	SP1/3/5(s)	SP1/3/4(s)	Avg.(s)
b	1.8±1.3	2.6±1.2	2.7±1.3	1.5±0.8	1.9±0.9	2.1±1.1	2.3±1.0	1.6±1.0	1.5±0.6	1.8±1.0	2.0±1.1
d	2.2±1.0	2.3±1.0	2.3±1.2	2.0±0.9	2.4±1.3	1.9±1.2	2.3±1.1	1.8±1.2	1.8±0.9	2.0±0.9	2.1±1.1

*Groups a, b, c and d are the same as those in Table 2.

Chai et al., 2022) and the exclusion of similar texture labels. It should be noted that during the human-machine interaction, the number of categorization labels can be flexibly adjusted to apply to the requirements of different texture recognition tasks. However, the actual discrimination performance should be reassessed by retraining the optimal algorithm model (such as, YOLOv5l).

5 Implications, limitations and future work

This article validated that the accuracies of identifying object surface roughness and texture via machine vision are higher than those of human tactile recognition. Object roughness and texture recognition using machine vision enables human-machine haptic interaction. The key significance lies in that it provides a practicable alternative for users who are willing to recognize object roughness or texture through tactile feedback when visual feedback is unavailable or they cannot directly contact the target objects. From an application perspective, the model SSD300 could be preferably applied for feature discrimination of object roughness or texture due to its optimal detection capability. Meanwhile, other types of algorithms for object surface feature recognition, such as the hybrid model of SSD and ResNet (Yang et al., 2021), can also be adopted to test whether the classification accuracy could be improved.

One limitation of this study is that the identification accuracies of the current machine vision models may slightly reduce when the roughness/texture levels or categories are increased, despite the classification capabilities are sufficient for a wide range of human-machine haptic interaction tasks. Therefore, in light of practical application requirements, design of appropriate training datasets is crucial to the performances of roughness/texture identification models. Moreover, the actual human-machine haptic interaction should take the object feature (roughness/texture) classification and its tactile feedback into consideration. Therefore, in the follow-up study, a miniature depth camera, an optimal roughness/texture classification model, a custom-made electrical stimulator, and a multi-electrode array will be integrated with an intelligent robotic device for practical use. Based on this, a battery of application-relevant object roughness/texture discrimination experiments with varied difficulty levels (e.g., multiple roughness/texture categories, different manipulation time) on various scenarios (e.g., with/without visual feedback / virtual reality (VR) guidance) will be designed to comprehensively investigate the benefits of bidirectional haptic interaction.

Data availability statement

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

Ethics statement

The studies involving humans were approved by Ningbo Institute of Materials Technology and Engineering. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ZL: Data curation, Formal analysis, Methodology, Software, Writing—original draft, Investigation, Resources, Validation. HZ: Data curation, Formal analysis, Investigation, Writing—original draft. YL: Formal analysis, Validation, Writing—original draft, Data curation. JZ: Conceptualization, Validation, Writing—review & editing, Formal analysis, Software. GC: Conceptualization, Funding acquisition, Project administration, Writing—review & editing, Methodology, Supervision, Validation. GZ: Resources, Writing—review & editing, Supervision, Project administration.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported in part by the Zhejiang Provincial Key Research and Development Program of China (2023C03160), the Ningbo Young Innovative Talents Program of China (2022A-194-G), Zhejiang Province Elite Program Project (2023C03162), and the Natural Science Foundation of Ningbo (2021J034).

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.

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

- Achancraray, D., Izumi, S.-I., and Hayashibe, M. (2021). Visual-electrotactile stimulation feedback to improve immersive brain-computer interface based on hand motor imagery. *Comput. Intell. Neurosci.* 2021, 1–13. doi: 10.1155/2021/8832686
- Arakeri, T. J., Hasse, B. A., and Fuglevand, A. J. (2018). Object discrimination using electrotactile feedback. *J. Neural Eng.* 15:046007. doi: 10.1088/1741-2552/aabc9a
- Arcos-García, Á., Álvarez-García, J. A., and Soria-Morillo, L. M. (2018). Evaluation of deep neural networks for traffic sign detection systems. *Neurocomputing* 316, 332–344. doi: 10.1016/j.neucom.2018.08.009
- Bajcsy, R. (1973). “Computer description of textured surfaces,” in *Proceedings of the 3rd International Joint Conference on Artificial Intelligence (IJCAI'73)* (San Francisco, CA: Morgan Kaufmann Publishers Inc.), 572–579.
- Chai, G., Wang, H., Li, G., Sheng, X., and Zhu, X. (2022). Electrotactile feedback improves grip force control and enables object stiffness recognition while using a myoelectric hand. *IEEE Trans. Neural Syst. Rehabil. Eng.* 30, 1310–1320. doi: 10.1109/TNSRE.2022.3173329
- Chapman, C. E., Tremblay, F., Jiang, W., Belingard, L., and Meftah, E.-M. (2002). Central neural mechanisms contributing to the perception of tactile roughness. *Behav. Brain Res.* 135, 225–233. doi: 10.1016/S0166-4328(02)00168-7
- Chen, N., Xu, Z., Liu, Z., Chen, Y., Miao, Y., Li, Q., et al. (2022). Data augmentation and intelligent recognition in pavement texture using a deep learning. *IEEE Trans. Intellig. Transp. Syst.* 23, 25427–25436. doi: 10.1109/TITS.2022.3140586
- Chen, Y.-M., Chou, F.-I., Ho, W.-H., and Tsai, J.-T. (2021). Classifying microscopic images as acute lymphoblastic leukemia by resnet ensemble model and taguchi method. *BMC Bioinform.* 22, 1–21. doi: 10.1186/s12859-022-04558-5
- Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., and Vedaldi, A. (2014). “Describing textures in the wild,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3606–3613.
- Clemente, F., Valle, G., Controzzi, M., Strauss, I., Iberite, F., Stieglitz, T., et al. (2019). Intraneural sensory feedback restores grip force control and motor coordination while using a prosthetic hand. *J. Neural Eng.* 16, 026034. doi: 10.1088/1741-2552/ab059b
- Connor, C. E., Hsiao, S. S., Phillips, J. R., and Johnson, K. O. (1990). Tactile roughness: neural codes that account for psychophysical magnitude estimates. *J. Neurosci.* 10, 3823–3836. doi: 10.1523/JNEUROSCI.10-12-03823.1990
- Dhanasekar, B., and Ramamoorthy, B. (2010). Restoration of blurred images for surface roughness evaluation using machine vision. *Tribol. Int.* 43, 268–276. doi: 10.1016/j.triboint.2009.05.030
- Farina, D., Vujaklija, I., Braanemark, R., Bull, A. M., Dietl, H., Graitmann, B., et al. (2023). Toward higher-performance bionic limbs for wider clinical use. *Nature Biomed. Eng.* 7, 473–485. doi: 10.1038/s41551-021-00732-x
- George, J. A., Kluger, D. T., Davis, T. S., Wendelken, S. M., Okorokova, E., He, Q., et al. (2019). Biomimetic sensory feedback through peripheral nerve stimulation improves dexterous use of a bionic hand. *Sci. Robot.* 4:eaax2352. doi: 10.1126/scirobotics.aax2352
- Georgiou, T., Liu, Y., Chen, W., and Lew, M. (2020). A survey of traditional and deep learning-based feature descriptors for high dimensional data in computer vision. *Int. J. Multimedia Inform. Retri.* 9, 135–170. doi: 10.1007/s13735-019-00183-w
- Glenn-jocher and Sergiuwaxmann (2020). *Comprehensive Guide to Ultralytics yolov5*. Available online at: <https://docs.ultralytics.com/yolov5/> (accessed March 15, 2024).
- Hashmi, A. W., Mali, H. S., Meena, A., Hashmi, M. F., and Bokde, N. D. (2023). Surface characteristics measurement using computer vision: A review. *CMES-Comp. Model. Eng. Sci.* 135:21223. doi: 10.32604/cmesci.2023.021223
- Hassan, W., Abdulali, A., Abdullah, M., Ahn, S. C., and Jeon, S. (2017). Towards universal haptic library: library-based haptic texture assignment using image texture and perceptual space. *IEEE Trans. Haptics* 11, 291–303. doi: 10.1109/TOH.2017.2782279
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- Hu, H., and Song, A. (2022). Haptic texture rendering of 2d image based on adaptive fractional differential method. *Appl. Sci.* 12:12346. doi: 10.3390/app122312346
- Jabban, L., Dupan, S., Zhang, D., Ainsworth, B., Nazarpour, K., and Metcalfe, B. W. (2022). Sensory feedback for upper-limb prostheses: opportunities and barriers. *IEEE Trans. Neural Syst. Rehabil. Eng.* 30, 738–747. doi: 10.1109/TNSRE.2022.3159186
- Jeyapooan, T., and Murugan, M. (2013). Surface roughness classification using image processing. *Measurement* 46, 2065–2072. doi: 10.1016/j.measurement.2013.03.014
- Jiang, P., Ergu, D., Liu, F., Cai, Y., and Ma, B. (2022). A review of yolo algorithm developments. *Procedia Comput. Sci.* 199, 1066–1073. doi: 10.1016/j.procs.2022.01.135
- Johnson, K. O., and Hsiao, S. S. (1994). Evaluation of the relative roles of slowly and rapidly adapting afferent fibers in roughness perception. *Can. J. Physiol. Pharmacol.* 72, 488–497. doi: 10.1139/y94-072
- Lin, P.-H., and Smith, S. (2018). “A tactile feedback glove for reproducing realistic surface roughness and continual lateral stroking perception,” in *Haptics: Science, Technology, and Applications: 11th International Conference, EuroHaptics 2018, Pisa, Italy, June 13-16, 2018, Proceedings, Part II 11* (Pisa: Springer), 169–179.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). “SSD: Single shot multibox detector,” in *Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I 14* (Netherlands: Springer), 21–37.
- Luo, S., Angrick, M., Coogan, C., Candrea, D. N., Wyse-Sookoo, K., Shah, S., et al. (2023). Stable decoding from a speech BCI enables control for an individual with als without recalibration for 3 months. *Adv. Sci.* 10:e2304853. doi: 10.1002/adv.202304853
- Makin, T. R., Micera, S., and Miller, L. E. (2023). Neurocognitive and motor-control challenges for the realization of bionic augmentation. *Nature Biomed. Eng.* 7, 344–348. doi: 10.1038/s41551-022-00930-1
- Mukoya, E., Rimiru, R., Kimwele, M., Gakii, C., and Mugambi, G. (2023). Accelerating deep learning inference via layer truncation and transfer learning for fingerprint classification. *Concurr. Comp.: Pract. Exp.* 35:e7619. doi: 10.1002/cpe.7619
- O’Sullivan, L., Picinali, L., and Cawthorne, D. (2014). “A prototype interactive tactile display with auditory feedback,” in *Irish HCI Conference* (Dublin: Dublin City University).
- Page, D. M., George, J. A., Kluger, D. T., Duncan, C., Wendelken, S., Davis, T., et al. (2018). Motor control and sensory feedback enhance prosthesis embodiment and reduce phantom pain after long-term hand amputation. *Front. Hum. Neurosci.* 12:352. doi: 10.3389/fnhum.2018.00352
- Rosén, B., Ehrsson, H. H., Antfolk, C., Cipriani, C., Sebelius, F., and Lundborg, G. (2009). Referral of sensation to an advanced humanoid robotic hand prosthesis. *Scand. J. Plastic Reconst. Surg. Hand Surg.* 43, 260–266. doi: 10.3109/02844310903113107
- Samra, R., Zadeh, M. H., and Wang, D. (2011). Design of a tactile instrument to measure human roughness perception in a virtual environment. *IEEE Trans. Instrum. Meas.* 60, 3582–3591. doi: 10.1109/TIM.2011.2161149
- Sensinger, J. W., and Dosen, S. (2020). A review of sensory feedback in upper-limb prostheses from the perspective of human motor control. *Front. Neurosci.* 14:345. doi: 10.3389/fnins.2020.00345
- Sigrist, R., Rauter, G., Riener, R., and Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychon. Bulletin. Rev.* 20:21–53. doi: 10.3758/s13423-012-0333-8
- Sklar, A. E., and Sarter, N. B. (1999). Good vibrations: tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. *Hum. Fact.* 41, 543–552. doi: 10.1518/00187209979656716
- Svensson, P., Wijk, U., Björkman, A., and Antfolk, C. (2017). A review of invasive and non-invasive sensory feedback in upper limb prostheses. *Expert Rev. Med. Devices* 14, 439–447. doi: 10.1080/17434440.2017.1332989
- Tong, Q., Wei, W., Zhang, Y., Xiao, J., and Wang, D. (2023). Survey on hand-based haptic interaction for virtual reality. *IEEE Trans. Haptics* 16, 154–170. doi: 10.1109/TOH.2023.3266199
- Way, T. P., and Barner, K. E. (1997). Automatic visual to tactile translation. I. human factors, access methods and image manipulation. *IEEE Trans. Rehabil. Eng.* 5, 81–94. doi: 10.1109/86.559353
- Weber, D., and Matsiko, A. (2023). Assistive robotics should seamlessly integrate humans and robots. *Sci. Robot.* 8:eadl0014. doi: 10.1126/scirobotics.adl0014
- Yang, D., Gu, Y., Thakor, N. V., and Liu, H. (2019). Improving the functionality, robustness, and adaptability of myoelectric control for dexterous motion restoration. *Exp. Brain Res.* 237, 291–311. doi: 10.1007/s00221-018-5441-x
- Yang, Y., Wang, H., Jiang, D., and Hu, Z. (2021). Surface detection of solid wood defects based on ssd improved with resnet. *Forests* 12:1419. doi: 10.3390/f12101419
- Zhao, Z.-Q., Zheng, P., Xu, S., and Wu, X. (2019). Object detection with deep learning: a review. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 3212–3232. doi: 10.1109/TNNLS.2018.2876865