

3D Touch Force Estimation from Capacitive Images

Jinyang Yu

yy-yu20@mails.tsinghua.edu.cn

Department of Automation, BNRist, Tsinghua University
Beijing, China

Jianjiang Feng*

jfeng@tsinghua.edu.cn

Department of Automation, BNRist, Tsinghua University
Beijing, China

Chentao Li

lct23@mails.tsinghua.edu.cn

Department of Automation, BNRist, Tsinghua University
Beijing, China

Jie Zhou

jzhou@tsinghua.edu.cn

Department of Automation, BNRist, Tsinghua University
Beijing, China

Abstract

Force plays a significant role in our daily interactions with the surrounding environment. However, most touchscreen interactions utilize only touch location, neglecting the potential for incorporating touch force as an additional input modality due to the difficulty in accurately estimating touch force without extra sensors. In this study, we propose a method for estimating three-dimensional relative touch forces, including pressure perpendicular to the touchscreen surface and shear force parallel to the surface, using two raw capacitive images. We collected two datasets comprising raw capacitive data from a touchscreen and corresponding forces measured by a triaxial force sensor. In the first dataset, participants performed press actions to apply pressure, while in the second dataset, they performed push actions to apply shear force to the touchscreen. Empirical experiments demonstrated that our proposed method outperformed existing force estimation methods, achieving mean absolute errors of 0.41 N, 0.44 N, and 0.40 N in the lateral, longitudinal, and vertical directions, respectively. Additionally, we conducted a user study with four tasks to assess the performance of our method in real-world scenarios, encompassing force gestures, pressure control, shear force control, and object manipulation using a combination of conventional touch input and force input. Comparisons with other methods demonstrated the superiority of our approach.

CCS Concepts

• **Human-centered computing** → **Interaction techniques**.

Keywords

touch input, force input, 3d object manipulation

ACM Reference Format:

Jinyang Yu, Chentao Li, Jianjiang Feng, and Jie Zhou. 2025. 3D Touch Force Estimation from Capacitive Images. In *30th International Conference on Intelligent User Interfaces (IUI '25)*, March 24–27, 2025, Cagliari, Italy. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3708359.3712123>

*Corresponding author



This work is licensed under a Creative Commons Attribution 4.0 International License. *IUI '25, Cagliari, Italy*

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1306-4/25/03

<https://doi.org/10.1145/3708359.3712123>

1 Introduction

In recent years, touch input has emerged as a ubiquitous input method for mobile devices owing to its simplicity and intuitiveness. With the advent of increased computing power, mobile devices now support demanding applications such as computer-aided design (CAD) and 3D video games. However, conventional touch input is limited to capturing 2D touch points, which are insufficient for complex task completion. To address this, research efforts have sought to expand the input space by leveraging information from interacting hands and fingers, including gesture [15, 16, 36, 37], finger angle [20, 39, 48, 52, 54], hand pose [1, 8], touch force [5, 25, 43], contacting hand region [18, 27, 31, 33, 45], etc.

Among these modalities, touch force has attracted considerable attention from both academia and industry. This stems from the inherent human tendency to exert force when fingers contact surfaces, which reflects the user's interactive intent. Previous work primarily employed external sensors or attachments to measure touch force, such as cameras [11, 13, 49], speaker and microphone [41], electromyographic (EMG) sensors [3], pressure sensors [22, 24, 35], joysticks [17, 51], fingerprint sensors [32, 53], etc. However, the introduction of additional sensors typically increases complexity and cost. To enhance usability, other research efforts have focused on estimating force using common built-in sensors, such as barometers [42], accelerometers [23, 30] and capacitive sensors [5, 6, 43].

Although most touchscreens only utilize the 2D positions of touch points as input, the underlying capacitive sensors capture low-resolution images that contain information about the contacting object. However, capacitive images do not directly encode touch force information because force variations do not typically affect the capacitive properties of the object [43]. Despite the limitation, fingertip tissue is soft and easily deformed when pressed against a rigid surface. This deformation alters the signal detected by the capacitive sensor, resulting in different contact area sizes and shapes under different forces (see Figure 1). Recognizing this relationship, researchers have explored various methods to estimate touch force from capacitive images. Boring et al. [6] proposed a Fat Thumb technique that simulates pressure based on the contact size of the thumb. Heo and Lee [25] estimated shear force based on the movement of the contact area. Arif and Stuerzlinger [2] combined contact time and touch point movement to distinguish between two pressure levels. Quinn et al. [43] used a sequence of capacitive images to distinguish between force touch and other touch gestures. Bocek et al. [5] trained a deep neural network (DNN) to estimate touch

pressure. A common challenge faced by these existing force estimation algorithms is the limited precision caused by variations in finger sizes and shapes. As a result, most of these algorithms are unable to deliver accurate continuous input. Furthermore, pressure alone provides only one additional degree of freedom (DOF), which is insufficient for complex 3D manipulation tasks.

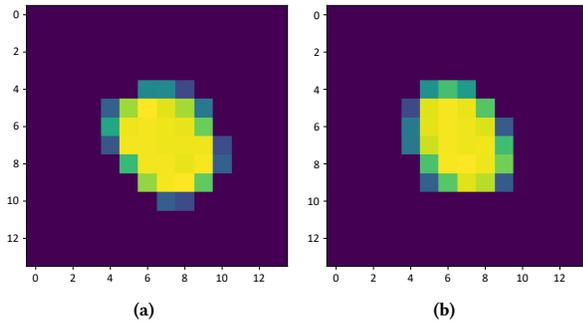


Figure 1: Two capacitive images of a finger touching a touchscreen with different forces: (a) under normal force; (b) under shear force.

Therefore, this paper proposes to estimate 3D relative forces from two capacitive images to support continuous 3D force interactions. The joint estimation of pressure and shear forces enhances accuracy by leveraging the inherent relationship between these two force components. To collect the necessary data, two tasks were designed in which participants applied normal and shear forces on a touchscreen while capacitive images and corresponding forces were recorded. A convolutional neural network (CNN) was employed to process two input capacitive images and estimate the 3D force differences. Experimental results demonstrated the superiority of the proposed force estimation method compared to existing absolute force estimation and indirect force estimation methods. The network model was deployed on a commodity smartphone, and a user study with four tasks was conducted to evaluate its performance in real-world scenarios, including force gestures, pressure control, shear force control, and object manipulation using a combination of traditional touch input and force input. The results confirmed the effectiveness of our approach. To promote reproducibility and enable further advancements in touch force estimation, we have open-sourced the code and data associated with this study: <https://github.com/radiance90/3DTouchForce>. Main contributions of this work:

- We propose to estimate 3D relative forces from two capacitive images, enhancing accuracy and robustness while providing additional DOFs for interactions.
- We train a lightweight CNN model to estimate both pressure and shear forces, achieving improved state-of-the-art accuracy in pressure estimation using capacitive images.
- We deploy the network model on a commodity smartphone and conduct a user study with four tasks to assess its performance in real-world scenarios.

2 Related Work

2.1 Touch Input

Touch input primarily relies on 2D touch point positions derived from low-resolution images captured by capacitive sensors [26]. Considerable effort has been dedicated to expanding the range of touch input modalities. Multi-touch gestures have undergone extensive research and have found practical applications in various input modalities due to their intuitive, flexible, and highly interactive user experience. Hancock et al. [15] introduced the Sticky Tools technique for 6DOF object manipulation using three-finger gestures. Martinet et al. [37] proposed DS3, a three-finger gesture based on the separation of translation and rotation. Liu et al. [36] extended the Sticky Tools technique to support 6DOF manipulation with two-finger gestures. One common drawback of gesture-based input is the gesture recognition failures, which can hamper the performance of gesture input. For instance, it is challenging to maintain a consistent distance between two fingers when performing a panning gesture, which can lead to the erroneous recognition of a pinching gesture [16]. Additionally, multi-touch gestures often occupy a significant portion of screen space, which can obstruct the display, especially on smaller screens. In contrast, utilizing force as an additional input modality involves only a single finger touch and eliminates the need for finger movements.

In recent years, finger angle has received significant attention as an additional input modality. Xiao et al. [52] utilized the shape of capacitive images to estimate yaw and pitch angles. Mayer et al. [39] employed a CNN to further reduce the errors associated with yaw and pitch angle estimation. Zaliva [54] leveraged multiple shape features, such as area, average intensity and centroids, to calculate three finger angles. He et al. [20] proposed TrackPose, a deep learning model with a self-attention module, to estimate yaw and pitch angles from a sequence of capacitive images. While finger angle estimation has shown promise, its accuracy is limited by the resolution of capacitive images.

With the advent of under-screen fingerprint sensors, fingerprints have been utilized to enhance the accuracy of finger angle estimation. Duan et al. [9] proposed a 2D-3D fingerprint matching algorithm for this purpose. He et al. [19] proposed a deep neural network with multi-task learning to estimate three finger angles directly from fingerprints. Duan et al. [10] further developed an algorithm to estimate 3D relative finger angles based on two fingerprint images, achieving state-of-the-art (SOTA) accuracy. Nonetheless, a major limitation of angle estimation using fingerprints is the limited availability of under-screen fingerprint sensors in commodity smartphones, which restricts the generalizability of this approach. The under-screen fingerprint sensors in commodity smartphones have limited sensing area, low capture rate and poor quality, affecting the performance and feasibility of finger angle input.

Another common limitation of using finger angle to augment touch input is the infeasibility of certain orientations [38], particularly in one-handed interaction scenarios. To address this issue, Roudaut et al. [44] utilized the kinematic traces on the touchscreen to distinguish 16 thumb gestures. Ullerich et al. [48] adopted a CNN model to detect the thumb’s pitch angle. However, the range of usable finger angles for input remains limited. In contrast, changes

in the forces applied to a touchscreen can be achieved without requiring finger movement or rotation. Thus, force input represents a potentially more natural and practical input modality that can complement conventional touch input.

Building upon finger angle, researchers have explored the use of capacitive images to predict hand pose. Ahuja et al. [1] proposed TouchPose, a multitask learning framework that predicts the depth image, 3D hand pose and touch classification. Choi et al. [8] introduced an algorithm that predicts hand pose by matching the capacitive images with reference hand poses. However, like multi-touch gestures, utilizing hand pose as an input method requires significant screen space, which can obstruct the view, particularly on smaller screens.

Furthermore, other modalities such as contacting hand positions [18, 27, 31, 33, 45] and finger identifications [12, 29, 34] have also been investigated to enrich the input vocabulary. However, these modalities are not specifically designed to provide continuous input for 3D manipulation tasks. By contrast, force offers a continuously changing vector, making it well-suited for continuous input.

2.2 Force Input

Using force as input provides more diverse and intuitive ways for human-computer interaction, enabling users to communicate and operate with computers and other intelligence devices in a more natural manner. Early studies on force input primarily focused on employing external sensors and attachments. Watanabe et al. [49] estimated touch force on a surface by detecting transmitted light on the fingernail. Becker et al. [3] measured finger touch forces using an electromyography armband. Grieve et al. [13] estimated 3D finger touch forces based on fingernail images. Fallahinia et al. [11] further estimated grasp forces of all fingers and the thumb based on fingernail images using a deep neural network. Ono et al. [41] proposed an active acoustic sensing technique to estimate touch force on the surface of an object. Heo and Lee [22] designed a mobile phone case-shaped frame with force sensing resistors attached to the bottom and side walls to measure normal and shear forces applied to the mobile device. Lee et al. [35] employed a similar device to investigate user controllability to reach and maintain shear force, considering the effects of hand pose and force direction. Harrison and Hudson [17] placed two analog joysticks underneath a touchscreen to enable 2D shear force input. Xiao et al. [51] implemented a similar structure on a smartwatch, expanding the range of available actions to include twist, tilt, and click. Huang et al. [28] proposed ShearSheet, a rubber-mounted transparent sheet on top of a touchscreen. The displacement of the sheet is recognized and utilized as shear input. Nakai et al. [40] applied a transparent gel layer on top of a touchscreen and measured shear force by analyzing the resulting deformation of the gel layer. These researches demonstrated the effectiveness of using force as an additional input modality. However, the use of external sensors often necessitates users to wear or modify their devices, resulting in additional costs and usability concerns.

To address this issue, researchers have focused on measuring touch forces using built-in sensors in commodity mobile devices. When airtight waterproof devices are touched, the distorted surface alters the air pressure inside, subsequently affecting the built-in

barometer value. Takada et al. [47] investigated the relationship between the sensor value and touch positions or forces, proposing BaroTouch, a technique that utilizes a waterproof mobile device's built-in barometer to measure touch force. Quinn [42] further refined this technique by developing a physical model for pressure equalization in devices, enabling high-resolution continuous tracking of user-independent touch forces. Although this approach demonstrates satisfactory performance under small forces, the sensing range is limited due to internal pressure saturation at the surface's distortion limit. In addition, accelerometers are also employed to detect touch force. Heo and Lee [23] proposed a technique to distinguish between gentle taps and forceful taps using acceleration in the z-axis (perpendicular to the device's surface). Hwang et al. [30] argued that the amount of vibration absorbed by the user's hand depends on the pressure applied to the device. Therefore, the authors utilized an accelerometer to measure the spatial displacement generated when the internal vibration motor vibrates and used this information to estimate the amount of pressure on the mobile device. However, due to accuracy limitations, accelerometer-based approaches can only distinguish discrete force levels. With the rapid advancement of under-screen fingerprint sensing technology, it has become possible to capture fingerprint images when a finger presses on the screen. Consequently, fingerprint images can be utilized as a novel input modality beyond user authentication. Fingerprint images contain abundant information about a finger, including clear ridges, minutiae, and touch area, among others. Analyzing the deformation of fingerprints under normal and shear forces enables estimation of the applied forces. Kurita et al. [32] proposed dividing a fingerprint into four quadrants and utilizing the ratios of their areas to estimate shear force. Yu et al. [53] introduced the PrintShear technique, which estimates shear forces by extracting lateral, longitudinal, and rotational deformations from fingerprint images. Research on shear force estimation based on fingerprint deformation has been particularly influential in guiding our investigation into the relationship between touch force and finger deformation. However, the limited sensing area, low capture rate, and poor image quality of under-screen fingerprint technology significantly impact the feasibility and performance of fingerprint-based force input. Moreover, the restricted availability of under-screen fingerprint sensors in commonly used smartphones limits the generalizability of fingerprint-based interaction techniques.

A touchscreen functions as a capacitive sensor capable of capturing low-resolution images of finger touch, which can be utilized for touch force estimation. Boring et al. [6] introduced the Fat Thumb interaction technique, which utilizes the contact size of the thumb as a form of simulated pressure. The mode of action is determined by the contact size, while the thumb's movement performs the manipulation. In a map navigation task under a one-handed interaction condition, this technique demonstrated superior performance compared to three existing techniques. While contact size can simulate force levels, it does not provide sufficient information for accurate force estimation due to variations in finger sizes. In contrast, capacitive images offer richer data, including size, shape, and deformation, which hold greater potential for precise force estimation. Heo and Lee [25] proposed an indirect method for estimating shear force by analyzing the movement of contact areas. To generate shear force, a normal force must be applied to the surface. The size of the contact

area reflects the applied normal force and is utilized to detect shear force events. When shear force is applied, the center of the contact area undergoes slight changes, and the displacement is roughly proportional to the magnitude of the shear force. An experiment was conducted to assess the feasibility of this method, and the results indicated that most errors were attributed to failures in detecting shear force events. Therefore, relying on the size of the contact area to detect shear force events is relatively unreliable, as the contact area is also influenced by finger pose and finger location. Quinn et al. [43] employed a sequence of capacitive images to differentiate between force touch and other touch gestures. Each capacitive image underwent feature extraction using a CNN. The extracted features were then processed by a recurrent layer to generate probabilities for different gesture classes. The network demonstrated the ability to distinguish six gestures, such as tap, long press, deep press, flick, and drag. Considering that tap, press, and scroll are the most common gestures on mobile devices, the authors decided to combine long press and deep press. The neural network was integrated into a heuristic classification pipeline to expedite the recognition of press gestures. Rather than introducing a new interaction modality, this work focuses on enhancing the user experience of long press interactions by accelerating them with force-induced deep press in a unified gesture. This approach maintains the effectiveness of long press gestures while providing a more natural and direct interaction. However, the study utilized capacitive image sequences solely for gesture classification. In contrast, our work focuses on the continuous estimation of touch forces to provide additional input dimensions. Boeck et al. [5] trained a deep neural network (DNN) to estimate touch pressure. To establish a baseline performance, well-established machine learning models such as k-nearest neighbors (KNN), decision tree (DT), support vector machine (SVM), and random forest (RF) were employed to estimate pressure. Although the proposed DNN outperformed all baseline methods, the accuracy of pressure estimation was insufficient for continuous force interaction. Researchers have made efforts to estimate force from capacitive images; however, the current lack of continuous force input support is attributed to the limitations in accuracy. Moreover, existing methods for estimating force based on capacitive images predominantly emphasize pressure, neglecting research on shear force estimation. This limitation restricts the extent of freedom that force input can offer.

3 Data Collection

The objective of this study is to estimate relative pressure and shear force based on capacitive images from a touchscreen. Due to the complex relationship between capacitive images and these forces, we employed a deep learning approach, which necessitates collecting paired capacitive images and corresponding ground-truth 3D force data. Since no existing dataset meets these requirements, we developed a specialized data collection apparatus and two data collection tasks to capture sequential capacitive images and synchronized 3D force values.

3.1 Apparatus

Capacitive images were acquired using a Realme C11(2021) smartphone¹, featuring a 6.5-inch main screen with a resolution of 1600 × 720 pixels. The capacitive sensor is 160 mm in length and 75 mm in width controlled by an ICNT8962 chip from Chipone Technology². It is capable of capturing 32 × 18 pixels capacitive images at 50Hz. The smartphone operates on Android 11, and the kernel has been modified to enable real-time access to capacitive images.

3D forces were measured using an Arizon AR-3N20S triaxial force sensor³. The sensor has a measurement range of 20 N for all three axes, with an error of less than 0.001%. The force sensor has a capture rate of 1500Hz. The smartphone was mounted in a 3D printed rigid plastic case that was securely attached to the force sensor (see Figure 2). The design of the case was based on the dimensions of the phone to prevent any displacement caused by shear force. Additionally, the force sensor was re-calibrated to account for the mass of the phone and the case. We validated the consistency of sensor measurements by placing weights at different positions on the screen. Both the phone and the force sensor were connected to the same laptop to ensure data synchronization.

Figure 3 shows the program we developed to collect and display real-time synchronized data. The user interface consists of three columns. The first column displays the normalized capacitive image and the maximum capacitive value within the image. The red dot indicates the centroid of the contract area. The second column provides a visualization of the 3D forces, including the raw force values. Lateral and longitudinal shear forces are represented by blue and green arrows, respectively, while the magnitude of the pressure is depicted by a red arrow. The third column consists of control buttons and text boxes for information recording.

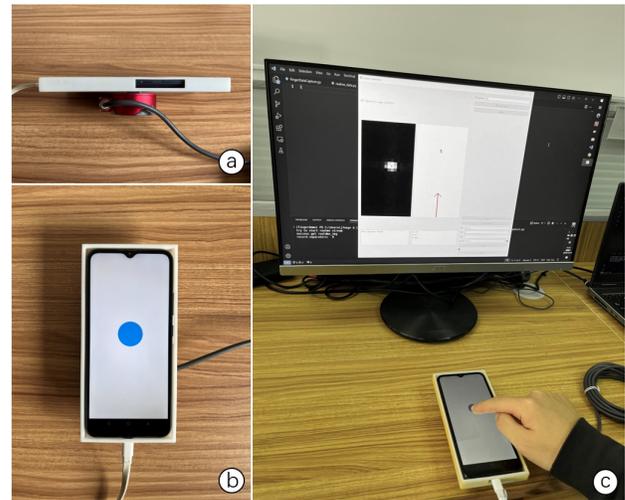


Figure 2: (a)-(b) Side and top views of the data collection apparatus. (c) The setup for the data collection tasks.

¹<https://www.realme.com/eu/realme-c11-2021/specs>

²<http://www.chiponeic.com/en/TT/229>

³<http://www.arizon-tech.com>



Figure 3: The data collection program consisting of the normalized capacitive image (A) with the maximum capacitive value (B), visualization of the 3D forces (C) with raw force values (D), and the control buttons (E).

3.2 Participants

For the data collection phase, we recruited 18 participants (15 males, 3 females) between the ages of 20 and 59 ($M=28.6$, $SD=11.6$). All participants were right-handed and had prior experience using touchscreen devices. Prior to their involvement, informed consent was obtained from each participant, and their personal information was treated confidentially. None of the participants had any medical conditions or upper limb injuries that would hinder their ability to complete the data collection tasks.

3.3 Task

Considering that pressure and shear force require distinct finger actions, we designed two tasks for each participant. In each task, participants were required to touch randomly generated blue circles on the touchscreen and perform either press or push actions. In the first task, participants were instructed to apply multiple presses on the touchscreen using varying forces. They were specifically asked to apply only normal force, which is perpendicular to the touchscreen, while attempting to minimize shear force. In the second task, participants were tasked with performing pushing actions in various forces and directions. To ensure accurate data collection, participants were instructed to avoid sliding their fingers on the touchscreen during both tasks.

3.4 Procedure

Upon welcoming the participants, we introduced the study’s objectives and demonstrated the apparatus and tasks. Participants completed a consent form and provided basic information, including age, gender, and dominant hand. The force sensor was positioned near the edge of a flat table, and the phone was aligned parallel to the table surface. The whole device was fixed firmly on the table to

avoid movement under pressure or shear forces. Participants were comfortably seated in front of the desk, ensuring ease of data collection. They were instructed to interact with the touchscreen using their typical finger angle, maintaining consistency throughout the tasks (see Figure 2c). Furthermore, participants were advised to refrain from touching anything other than the touchscreen.

Prior to each task, participants were given sufficient time to familiarize themselves with the task and its requirements. As the thumb and index finger are commonly used for touchscreen interactions, participants were asked to complete the tasks using these two fingers. To enhance dataset diversity, participants were asked to repeat the tasks using a lower finger pitch angle, specifically utilizing the finger pad region. After the data collection process, only frames that had a positive pressure value were retained. This step was taken to remove frames where there was no finger touching the touchscreen. In total, we collected 144 sequences (18 participants \times 2 fingers \times 2 finger angles \times 2 tasks), comprising 85,055 samples. Table 1 provides basic statistics for the two datasets. Although shear forces were still observed in the first dataset, they can be considered negligible in comparison to the normal force.

After the initial data collection, the dataset was subsequently organized into pairs. Each pair consisted of two capacitive images and their corresponding 3D relative forces. It’s important to note that only data pairs from the same sequence were considered, as the relative force measurements between different sequences lack significance for the purpose of interaction. Furthermore, pairs with a time difference greater than one second (50 frames) were discarded to ensure that the data came from continuous actions. A substantial total of 3,555,367 pairs, comprising capacitive images and their corresponding 3D relative forces, were successfully gathered. The distribution of the relative forces are illustrated in Figure 4. The two datasets were merged and subsequently divided randomly into a training set and a testing set, employing a 14:4 participant-wise split. This ensured a balanced distribution of data for training and evaluation purposes.

4 Method

This paper presents a novel approach for estimating 3D relative forces by employing a siamese framework, which utilizes two neural networks with identical architecture and shared weights to extract features from pairs of capacitive images. These extracted features are subsequently concatenated to estimate the 3D relative forces. Firstly, we provide a comprehensive explanation of the data preprocessing steps applied to the input capacitive images. Next, we delve into the structure of the siamese framework employed in this study. Finally, we elucidate the training strategy that was adopted for this framework.

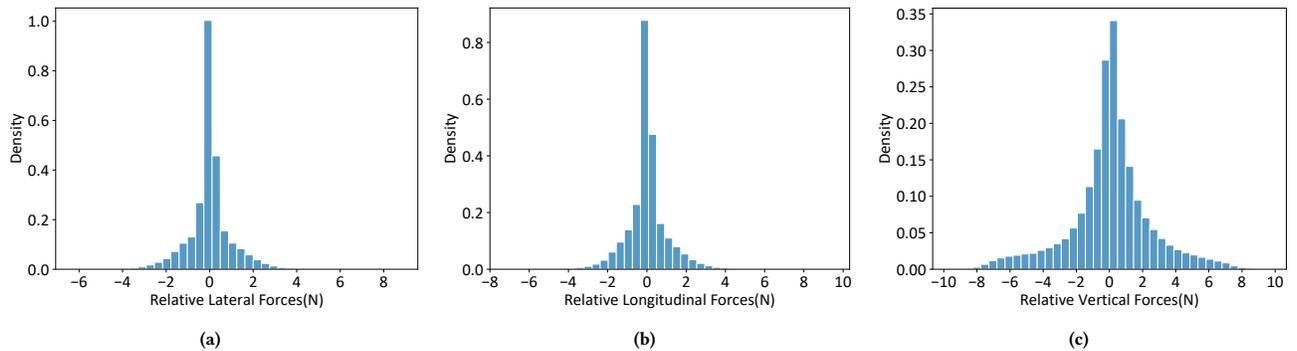
4.1 Data Preprocessing

To begin with, data normalization was performed on the two input images by dividing the maximum value of the first image. This approach, as opposed to separate normalization where both images are normalized to the range of 0 to 1, preserves the relative value changes of the data.

In order to enhance the position robustness of the network model, a center alignment strategy was adopted. The centroid of the contact

Table 1: 3D force statistics for the two datasets: the first dataset features press actions and the second dataset features push actions.

| | Press Action | | | Push Action | | |
|----------|--------------|--------------|----------|-------------|--------------|----------|
| | Lateral | Longitudinal | Vertical | Lateral | Longitudinal | Vertical |
| Min (N) | -1.32 | -2.48 | 0.01 | -5.50 | -5.19 | 0.01 |
| Max (N) | 1.04 | 1.20 | 11.10 | 3.76 | 4.63 | 11.46 |
| Mean (N) | -0.08 | -0.13 | 4.06 | 0.02 | 0.01 | 4.04 |
| SD (N) | 0.27 | 0.37 | 2.64 | 0.97 | 1.03 | 2.00 |

**Figure 4: Relative force distribution in the dataset: (a) lateral; (b) longitudinal; (c) vertical.**

area was obtained as the weighted mean of the pixel coordinates within the contact area. Subsequently, the image was shifted to align the centroid with the center coordinates of the image. Finally, the image was padded to a size of 32×32 pixels in order to better suit the network model.

4.2 Model Design

We propose to use a siamese network to estimate 3D relative forces (see Figure 5). Kurita et al. [32] and Yu et al. [53] demonstrated that fingers deform when pressed and pushed along a surface, and this deformation can be leveraged to estimate shear forces. When a network model is provided with a pair of capacitive images to estimate relative forces, it can extract deformation features in addition to shape and size features. In contrast, only shape and size features, among others, are available when estimating absolute forces. Furthermore, our pilot study on pressure estimation using contact size revealed that higher accuracy can be achieved by utilizing changes in contact size across two images rather than relying on the contact size from a single image. Although a time-series model with more input images could potentially improve accuracy, it introduces delays in the initial frames when a finger first contacts the touchscreen [20]. Additionally, processing more images increases computational demands and energy consumption. The primary advantage of utilizing a siamese network lies in its ability to learn discriminative features from input pairs. By sharing weights and architecture between two identical subnetworks, the network can acquire a robust representation of each input [7]. The network consists of a feature extraction module and a force estimation module.

The feature extraction module adopts a modified ResNet-18 structure [21], taking into account the input size and the computing power of mobile devices. It comprises a convolution block with 32 kernels of size 3 and a stride of 1, followed by two residual blocks with 32 and 64 channels, respectively. The resulting feature maps are flattened and passed through a fully connected (FC) layer to generate a 128-dimensional feature. The two 128-dimensional features are concatenated and then fed into the force estimation module, which consists of two FC layers that output the 3D relative forces.

4.3 Model Training

For the training data, data augmentation was performed by horizontally flipping the image while reversing the sign of the lateral shear force. This approach helps diversify the finger yaw angles and simulate capacitive images of fingers from the left hand. The proposed network was implemented in PyTorch and trained on a single NVIDIA GeForce 3090. The network was optimized by minimizing the mean squared error (MSE) loss for all three force dimensions. We employed the Adam optimizer with an initial learning rate of 0.001 to update the network parameters. We adopted a ReduceLROnPlateau strategy that reduces the learning rate by a factor of 0.1 when the performance stops improving for 10 epochs, and the training procedure concludes after the learning rate has decayed three times.

5 Experiments

This section presents the performance of the proposed method and compares it with two baseline methods. As previously mentioned,

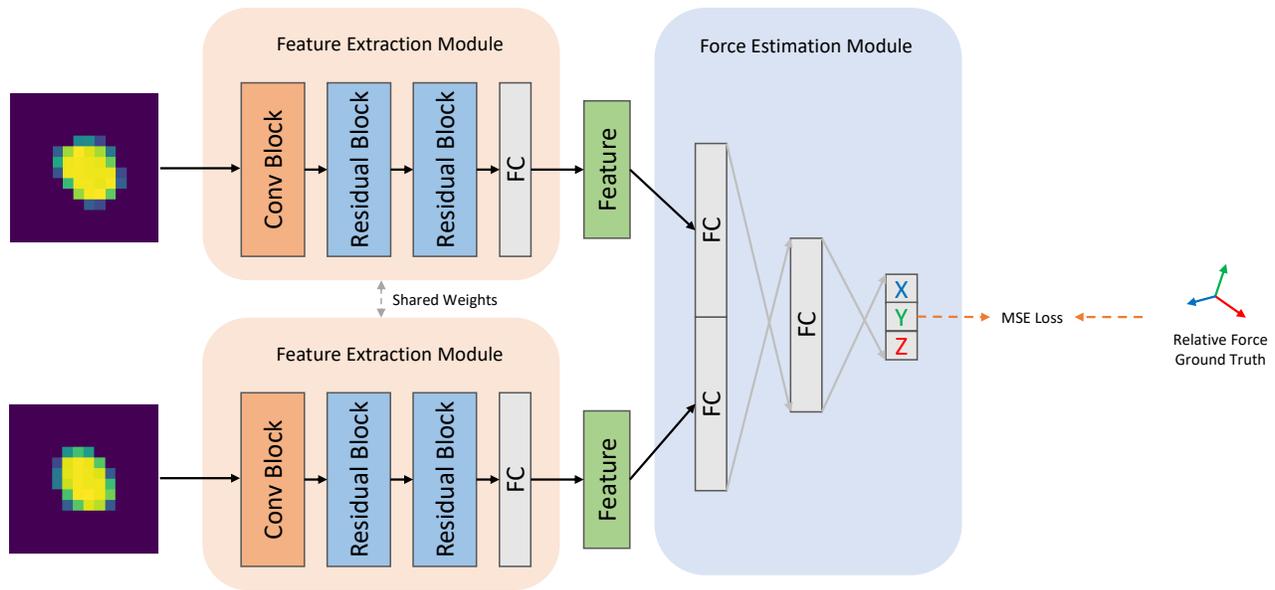


Figure 5: The proposed Siamese network consisting of a feature extraction module and a force estimation module.

the dataset was randomly divided into a training set and a testing set using a participant-wise split of 14:4. The evaluation is conducted on the testing dataset, which comprises 542,704 samples. The evaluation metrics used are mean absolute error (MAE), root mean squared error (RMSE), and standard deviation (SD).

5.1 Baseline Methods

In line with the studies by Heo et al. [25] and Bocek et al. [5], we implemented two baseline methods for evaluation. The first method involves an indirect estimation of forces through the assessment of size change and movement of the contact area, while the second method utilizes capacitive images for absolute force estimation.

5.1.1 Indirect Force Estimation. Heo et al. [25] proposed a method for indirectly estimating shear forces in multi-point shear force operations. The method is based on the hypothesis that the center of the contact area undergoes slight changes when shear force is applied, and the displacement is roughly proportional to the magnitude of the shear force. The authors also estimated pressure based on the size of the contact area. We re-implemented the methods and optimized the parameters using our training dataset. Initially, we employed the baseline approach described in [46] to extract the contact area from the input images. We enlarged the input images using cubic spline interpolation to obtain smoother boundaries, and then applied a 50% threshold value to determine the contact region. The contact size was calculated as the number of points within the contact region, while the centroid was obtained as the weighted mean of the pixel coordinates within that region. We fitted a linear regression model to estimate pressure changes using the difference in contact area size between two images. Additionally, two linear regression models were fitted using the movement of the centroid in lateral and longitudinal directions for shear force estimation.

5.1.2 Absolute Force Estimation. Bocek et al. [5] proposed a DNN for estimating the pressure applied to the touchscreen. The network consists of three convolutional layers and two fully connected layers. However, as the original model was designed solely for pressure estimation, the final fully connected layer produced only a single output. To adapt the model for 3D force estimation, we modified the last layer to output three values, corresponding to the three force dimensions. The model was then retrained using the data collected in Section 3, following the same training strategy as outlined in the original study.

5.2 Results and Discussion

The experimental results for both the proposed method and the baseline methods are presented in Table 2. The proposed approach for estimating relative force demonstrated the lowest values for RMSE, MAE, and SD across all three dimensions, indicating the highest accuracy and reliability compared to other methods. Specifically, the proposed method achieved MAEs of 0.41 N, 0.44 N, and 0.40 N for the three dimensions, representing reductions of 31.67%, 22.81%, and 39.39%, respectively, compared to the SOTA force estimation method by Bocek et al. [5]. Additionally, in terms of pressure estimation, the indirect method showed significantly higher errors and variations compared to the other two methods, suggesting that using contact size for pressure estimation is not sufficiently robust to account for differences in finger size. The force estimation error distributions are presented in Figure 6, showing that errors are predominantly concentrated in the low-error region when using the proposed method.

5.2.1 Analysis of 3D Force Joint Estimation. Additional experiments were conducted to examine the impact of joint 3D force estimation. In addition to estimating both pressure and shear forces with a single model, we implemented two separate models to estimate

Table 2: Quantitative comparison of 3D force estimation errors for different methods. Errors are reported in Newtons (N).

| Method | Lateral | | | Longitudinal | | | Vertical | | |
|---|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|
| | RMSE | MAE | SD | RMSE | MAE | SD | RMSE | MAE | SD |
| Indirect Estimation by Heo et al. [25] | 0.82 | 0.63 | 0.53 | 0.90 | 0.65 | 0.62 | 2.19 | 1.64 | 1.45 |
| Absolute Estimation by Bocek et al. [5] | 0.80 | 0.60 | 0.53 | 0.80 | 0.57 | 0.56 | 0.85 | 0.66 | 0.54 |
| Ours | 0.57 | 0.41 | 0.40 | 0.63 | 0.44 | 0.45 | 0.53 | 0.40 | 0.33 |

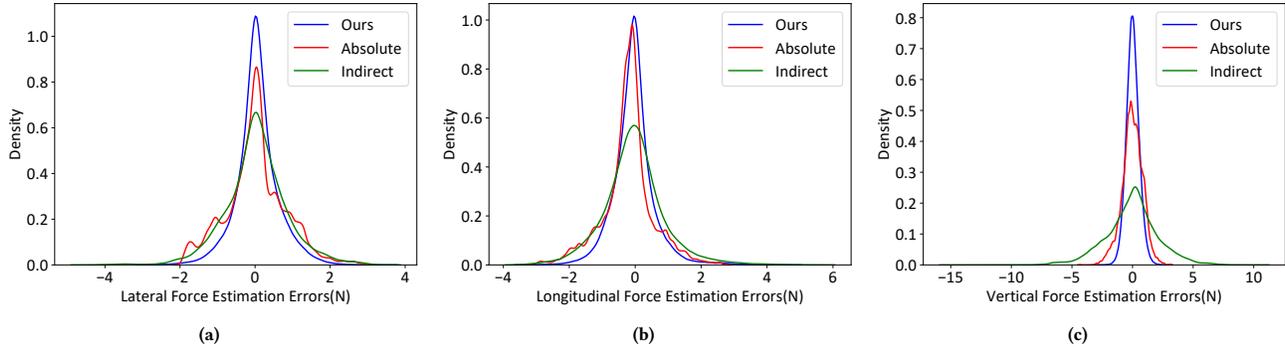


Figure 6: 3D force estimation error distributions: (a) lateral; (b) longitudinal; (c) vertical.

pressure and shear forces individually, modifying the number of outputs in the final fully connected layer accordingly. As shown in Table 3, the model with joint 3D force estimation outperformed the separate models across all three force dimensions. We attribute this improvement to the intrinsic relationship between pressure and shear forces. When a finger pushes on a touchscreen without sliding, the static friction opposes the shear force. The maximum static friction is proportional to the pressure applied on the surface. This relationship can serve as a constraint during estimation. For example, it is impossible to exert a large shear force while applying minimal pressure.

5.2.2 Performance on Pressure Dataset. Unlike shear force, pressure can be applied and utilized independently. Most algorithms and applications focus on leveraging pressure as an additional input modality. Therefore, in this experiment, we trained and evaluated the proposed method and baseline methods using the pressure dataset. It is important to note that the original pressure estimation model by Bocek et al. [5] was retrained without any modifications to its network structure. As shown in Table 4, the proposed approach achieved the lowest values for RMSE, MAE, and SD, recording 0.64 N, 0.49 N, and 0.41 N, respectively. Notably, the proposed method demonstrated a 40.24% reduction in MAE compared to Bocek et al.’s force estimation method. This indicates that, in addition to incorporating shear modalities, the proposed method enhances accuracy in pressure estimation tasks.

6 User Study

To assess the feasibility and performance of various 3D force estimation methods in a real-world setting, we implemented all methods

on a commodity smartphone and carried out a user study comprising the following four tasks:

- Force gestures performing task.
- Pressure control task.
- Shear force control task.
- 4 DOF object manipulation task using a combination of conventional touch input and force input.

These tasks also showcased several potential use cases where continuous 3D force can be employed as input.

6.1 Apparatus and Model Deployment

The Realme C11 (2021) smartphone, which was employed for data collection, was also utilized in the user study. The applications for all tasks were developed using Android Studio⁴. We implemented the aforementioned three force estimation methods on the smartphone. To enable the execution of network models on the device, we initially generated dedicated OpenVINO⁵ library files tailored for the Realme smartphone’s armeabi-v7a architecture. Subsequently, we integrated these files into the Android applications required for each task. Furthermore, we utilized the OpenVINO Toolkit to convert the network models and imported the converted models for inference.

Unlike the data collection task where measured pressures are employed to detect finger touch, the user study requires a finger detection algorithm for all applications involved. To filter the raw capacitive image, a threshold value of 200 was utilized. Any point in the capacitive image with a value below the threshold is assigned a value of zero. Finger touch is detected if there exist non-zero

⁴<https://developer.android.com/studio>

⁵<https://docs.openvino.ai/2024/home.html>

Table 3: Quantitative comparison of 3D force estimation errors using separate pressure and shear force estimation methods versus the joint 3D force estimation method. Errors are reported in Newtons (N).

| Method | Lateral | | | Longitudinal | | | Vertical | | |
|---------------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|
| | RMSE | MAE | SD | RMSE | MAE | SD | RMSE | MAE | SD |
| Separate Estimation | 0.62 | 0.43 | 0.41 | 0.70 | 0.49 | 0.48 | 0.59 | 0.46 | 0.37 |
| Joint Estimation | 0.57 | 0.41 | 0.40 | 0.63 | 0.44 | 0.45 | 0.53 | 0.40 | 0.33 |

Table 4: Quantitative comparison of pressure estimation errors for different methods. Errors are reported in Newtons (N).

| Method | RMSE | MAE | SD |
|-------------------------------------|-------------|-------------|-------------|
| Indirect Method by Heo et al. [25] | 2.37 | 1.79 | 1.55 |
| Absolute Method by Bocek et al. [5] | 1.09 | 0.82 | 0.72 |
| Ours | 0.64 | 0.49 | 0.41 |

values in the filtered image. This approach effectively differentiates between touch and non-touch, particularly in scenarios where a finger is in close proximity but not directly touching the screen. Once the initial frame is identified, the relative forces between the initial frame and subsequent frames are estimated.

6.2 Participants

We recruited 12 participants (10 males, 2 female) aged between 20 and 33 ($M=22.92$, $SD=3.45$) who were not involved in the data collection process. One participant was left-handed, while the remaining participants were right-handed. All participants had previous experience with touchscreen devices, while two participants had experience with force input (Apple 3D Touch). Prior to their participation, we obtained informed consent from each participant. None of the participants had any medical conditions or upper limb injuries that could impede their ability to complete the user study.

6.3 Tasks

6.3.1 Task 1: Force gestures. Previous studies have explored the use of touch force as a gesture or operation mode switcher [6, 23, 24], though most focused exclusively on pressure gestures. In this study, we introduce six distinct force gestures: gentle press, hard press, push left, push right, push up, and push down. These gestures are categorized based on force changes, as shown in Figure 7. Each gesture starts when a finger makes contact with the screen and ends when the finger is lifted. We estimate and track the 3D relative forces between the first touch frame and each subsequent frame. To distinguish between traditional swipe gestures and force gestures, we employed a pressure threshold of 2 N. If the maximum pressure during the gesture is below 2 N, the gesture is categorized as a gentle press; otherwise, it is classified as a force gesture. Additionally, if the magnitude of the lateral and longitudinal shear forces remains consistently below 1 N throughout the gesture, it is classified as a hard press. If the maximum shear force exceeds this threshold, the gesture is then classified as a push gesture. The shear force threshold was determined by analyzing data from the press action

dataset, where less than 4% of samples exhibited a shear force above 1 N, with this percentage dropping to below 1% for lighter presses (pressure below 6 N). The direction of the push gesture is determined by the dominant shear force component—either lateral or longitudinal.

The aim of this task is to evaluate the recognition accuracy of force gestures based on the force estimated by various methods. Figure 8a displays a screenshot of the application used for this task. At the beginning of the task, a target gesture is displayed on the screen. After completing a gesture, the screen will flash either green or yellow, indicating whether the detected gesture was correct or incorrect. Subsequently, the next target gesture is presented on the screen. Both the target gestures and the performed gestures are recorded. Overall, a total of 2160 trials were conducted in this task, which involved 12 participants, 3 methods, 6 gestures, and 10 repetitions.

6.3.2 Task 2: Pressure control. The objective of this task is to assess the controllability and stability of continuous pressure control using different force estimation methods. The application employed for this purpose is depicted in Figure 8b. The top half of the screen features a progress bar that can be controlled by applying pressure. At 0 N pressure, the progress bar is empty, and it becomes full at 10 N. Two red markers are randomly positioned on the progress bar, maintaining a fixed distance of 10% of the bar’s length between them. To ensure a fair comparison among the methods, the marker positions are selected from the same pool of randomly generated positions. Participants are instructed to adjust the progress by modulating their pressing force. A trial is considered complete if the progress remains between the two markers for a continuous duration of 0.5 seconds. Once a trial is completed, the time taken is recorded, and the markers are relocated to new random positions. In total, this task consisted of 720 trials, involving 12 participants, 3 methods, and 20 repetitions.

6.3.3 Task 3: Shear force control. Similar to Task 2, the objective of this task is to assess the controllability and stability of continuous shear force control using different force estimation methods. The application utilized for this task is depicted in Figure 8b. The lower half of the screen is dedicated to this particular task, displaying a target circle with a radius of 50 pixels and a dot with a radius of 10 pixels. To ensure a fair comparison among the methods, the positions of the target circle are selected from the same pool of randomly generated positions. The maximum shear force required to move the dot to the boundary is 3 N in all directions. Participants are instructed to manipulate the small dot by adjusting their shear forces. A trial is considered successful if the small dot remains completely within the large circle continuously for a duration of

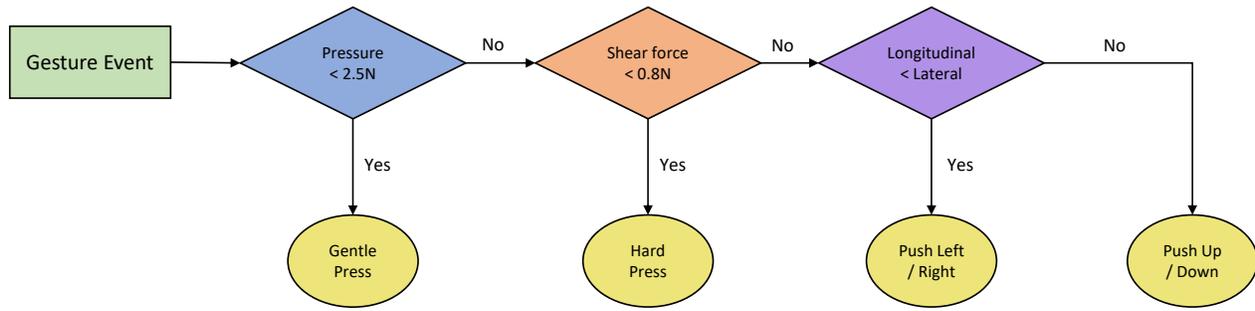


Figure 7: The force gesture determination process.

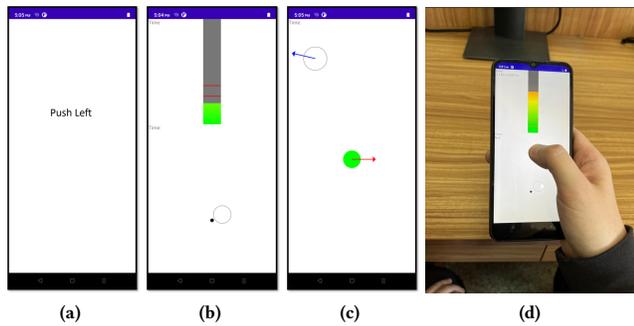


Figure 8: Applications developed for the user study: (a) a gesture performing task; (b) a pressure control task (upper half), and a shear force control task (lower half); (c) a 4DOF object manipulation task. (d) The setup for the user study.

0.5 seconds. A trial is deemed unsuccessful if it is not completed within a time frame of 15 seconds. Once a trial is completed, the time taken is recorded, the big circle moves to the next random position, and the small dot returns to its original position. In total, this task involved 720 trials, with 12 participants, 3 methods, and 20 repetitions.

6.3.4 *Task 4: 4 DOF object manipulation.* To make a new modality practical for commodity smartphones, it should be compatible with conventional touch input. Hence, the objective of this task is to assess the performance of various force estimation methods under this condition. Manipulating objects with 4DOF is a common task in daily mobile phone operations. For instance, users may need to pan, zoom, or rotate a map or a photo. The current approach to accomplish this task involves multi-finger gestures, which is not feasible for single-handed phone operation [6]. In this task, a new interaction technique for 4DOF is devised, as illustrated in Table 5. Conventional touch input is employed for panning, lateral shear force for rotation, longitudinal shear force for zooming, and pressure for switching operation modes. A pressure threshold of 2 N is used for switching between conventional touch input and force input. The application used for this task is presented in Figure 8c. A randomly sized target circle is positioned on the screen,

accompanied by a blue arrow indicating the target angle. To ensure a fair comparison among methods, the positions, size of the target circle, and angle of the arrow are selected from the same pool of randomly generated values. Participants are instructed to move a green pie to the target location, aligning its size and direction accordingly. A trial is considered successful if the location difference is less than 10 pixels, the radius difference is less than 10 pixels, and the angle difference is less than 5 degrees. A trial is considered unsuccessful if it exceeds a duration of 60 seconds. Once a trial is completed, the time taken is recorded, and the position, size and direction of the target circle is refreshed. Based on our pilot test, we found that the absolute approach was inadequate for this task due to its inaccurate and unstable shear force estimation. Consequently, we excluded the absolute force estimation method from consideration for this specific task. Overall, this task encompassed 240 trials, involving 12 participants, 2 methods, and 10 repetitions.

6.3.5 *Procedure.* The user study followed a procedure similar to the data collection process outlined in Section 3.4. Upon arrival, participants were briefed on the study’s objectives and given a demonstration of the apparatus and tasks. After providing informed consent, participants submitted basic demographic information, including age, gender, and dominant hand. They were seated comfortably in chairs, allowed to adopt their preferred posture, and instructed to interact with the phone using their preferred hand and preferred finger (see Figure 8d). As in the data collection phase, they were asked to maintain a consistent finger angle while interacting with the touchscreen. Before each task, participants were allotted sufficient time to familiarize themselves with the task requirements. To minimize potential bias from the learning effect, the order of input methods for each task was counterbalanced across participants using a Latin square design. After completing each task, participants provided subjective ratings through questionnaires. Additionally, they completed a NASA-TLX (raw TLX) questionnaire to assess perceived workload after finishing all tasks. Interviews were conducted at the end of the study to collect feedback on any issues encountered and participants’ opinions regarding each force estimation method. The study adhered strictly to ethical guidelines for research involving human subjects.

Table 5: Control mappings using a combined strategy of conventional input and force input.

| | Up | Down | Left | Right |
|--------------|---------|----------|----------------------|------------------|
| Conventional | Pan Up | Pan Down | Pan Left | Pan Right |
| Force | Zoom In | Zoom Out | Rotate Anticlockwise | Rotate Clockwise |

6.4 Performance Analysis

To assess the normality of the data across all tasks and subjective ratings, we conducted the Shapiro-Wilk test, which revealed that all time-related tasks (Tasks 2-4) violated the assumption of normality. Consequently, we used violin plots to visualize the actual distribution of the original data and applied non-parametric Wilcoxon signed-rank tests[50] with Benjamini-Hochberg corrections[4] to evaluate the significance of completion time differences between methods. For subjective ratings, we plotted the means and 95% confidence intervals (CIs) and performed paired t-tests with Benjamini-Hochberg corrections to determine statistical significance.

6.4.1 Task 1: Force gestures. The force gesture recognition rates are presented in Table 6. Among the different methods evaluated, the proposed 3D relative force estimation method demonstrated the highest accuracy, achieving an 95.97% overall recognition rate. This marks an improvement of 77.16% compared to the absolute force estimation method and a 43.37% improvement over the indirect force estimation method. The low recognition rate of the indirect approach for pressure gestures is consistent with the experimental results in Table 2, which we attribute to the method’s inability to robustly accommodate variations in finger size. The absolute approach exhibited the poorest recognition rate for shear gestures, supporting our hypothesis that estimating relative force from two images provides more accurate results. The confusion matrices for the three methods, shown in Figure 9, highlight that shear gestures were frequently misclassified as hard presses when using the absolute approach, reflecting its limited range in estimating shear force. Meanwhile, the indirect approach, which lacks the ability to detect force gestures, often misclassified shear gestures as gentle presses.

6.4.2 Task 2: Pressure control. The completion times for the pressure control task are illustrated in Figure 10a. The average completion times for the relative, absolute, and indirect approaches are 1.84 s, 2.43 s, and 2.60 s, respectively. The proposed approach demonstrated the highest performance, and the difference was statistically significant when compared to the absolute approach ($p < 0.001$) and the indirect approach ($p < 0.001$). These findings are consistent with the experimental results presented in Table 2 and the pressure gesture recognition results outlined in Table 6.

6.4.3 Task 3: Shear force control. The completion times for the shear force control task are presented in Figure 10b. Notably, only 23.75% of the trials using the absolute method were successful (completed within 15 seconds), consistent with its lower shear force gesture recognition rates shown in Table 6. As a result, the average completion time for the absolute method is not depicted in the figure. Although the indirect method exhibited similar shear force estimation errors to the absolute method (see Table 2), all participants successfully completed the task using the indirect approach.

This method estimates shear force based on movement alone; while the estimated values may deviate from the actual force, the linear model ensures consistent force changes between adjacent frames. The average completion times for the proposed method and the indirect approach were 2.07 s and 2.46 s, respectively, with the performance difference being statistically significant ($p < 0.001$).

6.4.4 Task 4: 4 DOF object manipulation. The completion times for the 4DOF object manipulation task are illustrated in Figure 11. The average completion times for the relative and indirect approaches are 18.62 s and 24.68 s, respectively. The proposed method achieved a 24.55% reduction in completion time compared to the indirect approach, with the difference being statistically significant ($p < 0.001$). In addition to the shear force control ability, consistency and accuracy in pressure estimation played a crucial role in task performance. Although pressure was primarily used as a mode switcher, incorrect mode classification led to discrepancies between the intended action and the participant’s actual action, potentially causing previously aligned dimensions to misalign.

6.4.5 Subjective Ratings. We obtained usability and preference ratings from the participants upon completing each task. The results are presented in Figure 12. Across all tasks, the proposed force estimation method received the highest ratings for both usability and user preference. In contrast, the absolute approach received the lowest ratings in both the force gesture and shear force control tasks, largely due to its poor performance in estimating shear force. Workload data, measured using the NASA-TLX, is presented in Figure 13. The proposed method consistently exhibited the lowest workload across all criteria, with a significantly lower total workload compared to the absolute approach ($p < 0.001$) and the indirect approach ($p < 0.01$). Participants reported higher frustration, increased effort, and lower performance with the absolute approach, particularly due to its failures in shear force operations.

6.5 Discussion

The user study revealed that the proposed 3D relative force estimation method achieved the highest overall performance, offering a balanced ability to estimate both pressure and shear forces. In contrast, the absolute approach performed poorly in both the force gesture and shear force control tasks, primarily due to its inability to estimate shear forces. This limitation arises because the absolute approach relies on the shape and capacitive intensities from a single capacitive image, without accounting for variations over time. The overall gesture recognition rate based on relative force estimation is 77.16% higher than that achieved with the absolute force estimation method, a difference more pronounced than the error reduction observed in the force estimation experiment. This suggests that low pressure and shear force estimation accuracies interfere with one another, further diminishing performance in

Table 6: Force gesture recognition rates for different force estimation methods.

| Method | Gentle Press | Hard Press | Push Up | Push Down | Push Left | Push Right | Overall |
|---|---------------|---------------|---------------|---------------|----------------|----------------|---------------|
| Indirect Estimation by Heo et al. [25] | 63.33% | 71.67% | 60.83% | 77.50% | 54.17% | 74.17% | 66.94% |
| Absolute Estimation by Bocek et al. [5] | 75.83% | 81.67% | 22.50% | 33.33% | 56.67% | 55.00% | 54.17% |
| Ours | 98.33% | 91.67% | 96.67% | 89.17% | 100.00% | 100.00% | 95.97% |

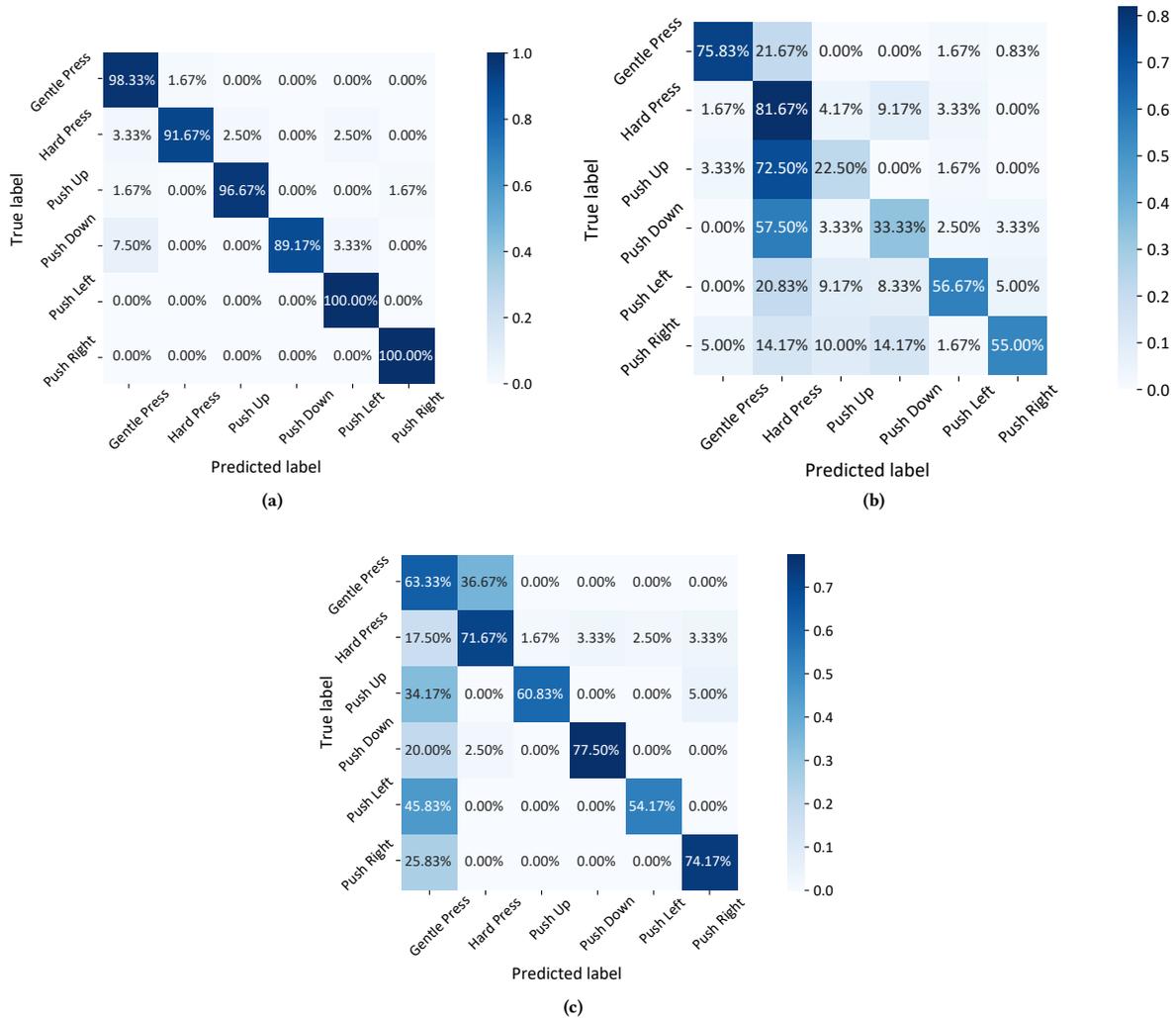


Figure 9: Confusion matrix for force gesture recognition based on (a) proposed relative force estimation; (b) absolute force estimation; (c) indirect force estimation.

real-world applications. Additionally, the improved shear force estimation facilitates continuous shear force control, as demonstrated in Task 3. Beyond task performance, the enhanced estimation accuracy also contributes to a better user experience by reducing effort and frustration levels.

Meanwhile, the indirect approach exhibited the weakest performance in the pressure control task, likely due to its inability to accommodate variations in finger size. However, the indirect

approach outperformed the absolute approach in the shear force control task, despite having slightly higher shear force estimation error rates (see Table 2). This can be attributed to the fact that the indirect approach relies solely on movement to estimate shear forces. While the estimated values may deviate from the actual force, the linear model ensures consistent force changes between adjacent frames, allowing participants to adjust their actions based

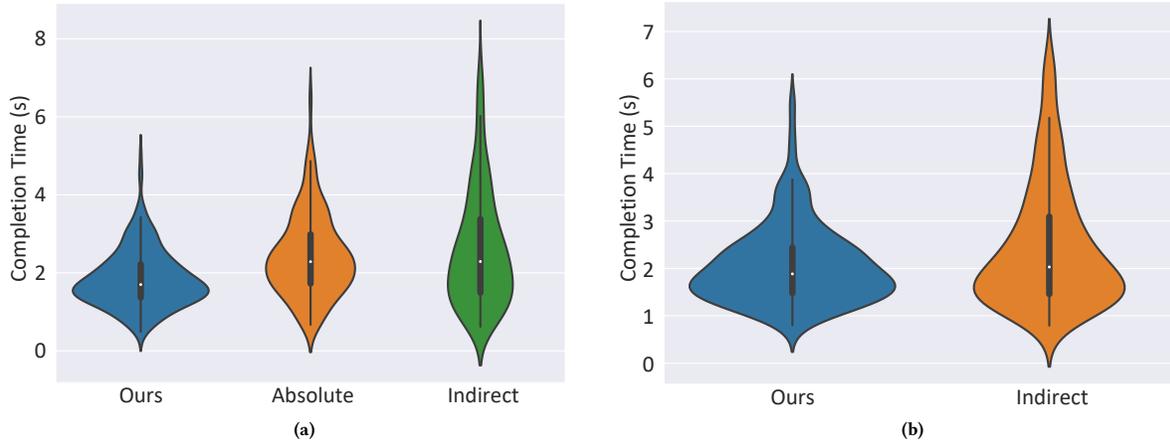


Figure 10: Completion times (in seconds) for (a) the pressure control task and (b) the shear force control task using different force estimation methods.



Figure 11: Completion times (in seconds) for the 4DOF object manipulation task using different force estimation methods.

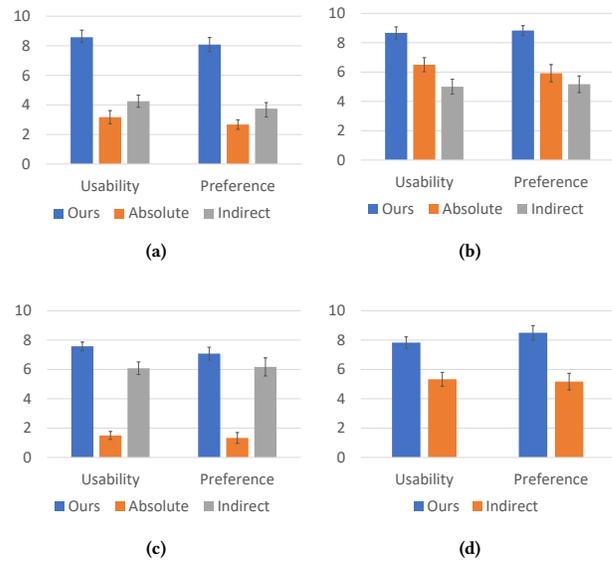


Figure 12: Subjective usability ratings and preference scores for each task of the user study: (a) force gesture task; (b) pressure control task; (c) shear force control task; (d) 4DOF object manipulation task. Error bars: 95% CI.

on visual feedback. In the absence of visual feedback, such as in force gesture tasks, performance declined significantly.

Additionally, some participants reported that the pressure range achievable with the indirect method was limited, requiring them to change the finger angle to achieve a broader range of pressure values. This limitation is due to the restricted deformation range of the fingertip, which limits the change in contact area size. By altering the finger angle, participants were able to increase the range of contact area sizes, thereby expanding the range of indirect pressure estimation.

7 Applications

We have identified three categories of force interactions enabled by the proposed method and implemented a series of applications to demonstrate the potential of 3D force interactions.

7.1 Force Gestures

Most force gestures introduced in previous studies [17, 28] can be replicated using our proposed method. Pressure gestures are the most intuitive and commonly used gestures. For instance, a firm press can be distinguished from a light tap, which can be employed to open pop-up menus (Figure 14a). The force-based firm press can serve as a replacement or complement to the widely used long press gesture, offering the advantage of requiring less time to execute[43]. With accurate pressure estimation, the number

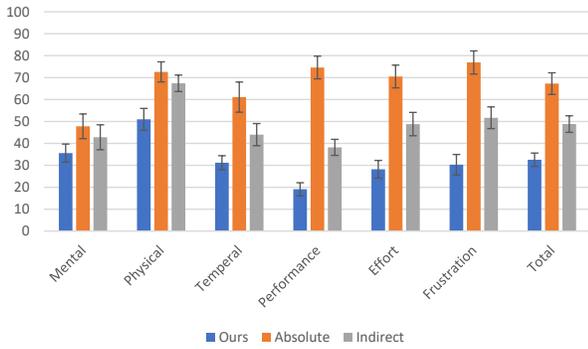


Figure 13: Workload measurement in NASA-TLX units. Error bars: 95% CI.

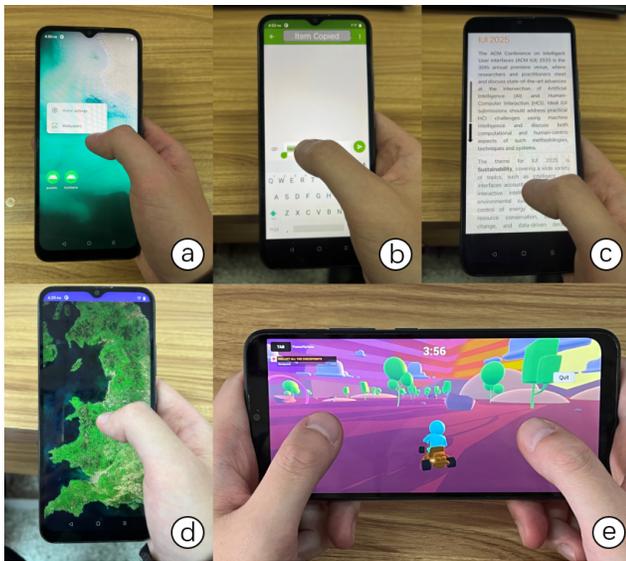


Figure 14: Applications of the proposed 3D relative force estimation method: (a) opening a pop-up menu with a firm press gesture; (b) copying text using an upward push gesture; (c) fast scrolling via rate-based shear force input; (d) map navigation through combined touch and force input; (e) controlling a kart in a game with two-finger force input.

of available menus can be expanded by incorporating different pressure levels. Additionally, 2D shear gestures offer the potential for a broader range of actions. Simple directional push actions, as discussed in Section 6.3.1, can be used for system-level commands, such as adjusting volume or brightness, or as application-specific shortcuts like copy, paste, and cut (Figure 14b).

7.2 Continuous Force Interactions

The proposed method enables fast and accurate force estimation, allowing for continuous force interactions. Force input can complement traditional touch input in two ways: by providing input with adjustable control-display (CD) gain and by offering extra degrees

of freedom (DOF). Conventional touch input typically operates at a 1:1 CD gain, which may be inefficient for tasks such as fine-grained manipulations or rapid scrolling. For example, when browsing content larger than the screen, users must repeatedly flick to scroll. With shear force-based high CD gain control, fast scrolling can be achieved with fewer flicks (Figure 14c). Conversely, shear force can be used with a low CD gain to enable precise object alignments.

Some applications require operations with more than two degrees of freedom. For instance, navigating a map typically involves panning, zooming, and rotating. Most map applications use two-finger gestures for zooming and rotating, which can be inconvenient when interacting with one hand. In this case, longitudinal and lateral shear forces can control zooming and rotation, respectively (Figure 14d).

Additionally, continuous force interactions can be combined with force gestures. For example, shear forces can navigate a pop-up menu triggered by a pressure gesture.

7.3 Multi-finger Interactions

Unlike under-screen force sensors that measure the total pressure applied to a touchscreen, the proposed method can estimate the forces applied by each individual finger. After detecting the fingers, the system processes the image patches containing each finger separately, allowing for the estimation of the touch force of each finger independently. These force estimates can then be used to control different DOFs or recognized as multi-finger force gestures. Figure 14e shows a kart racing game we implemented that can be controlled using two-finger shear force input. The shear force of the left thumb controls the camera angle, while the shear force of the right thumb controls the movement of the kart.

8 Limitations and Future Work

The proposed 3D relative force estimation method was evaluated through empirical experiments and a user study, demonstrating its feasibility and performance. However, this study has the following limitations:

The dataset collected in this study only included data from the thumb and index finger, with a limited range of finger angles and touch positions. In real-world scenarios, users interact with touchscreens using various fingers and finger angles at different positions, so a more diverse dataset would improve the generalizability of the proposed method. Additionally, the study had a limited number of participants for data collection and the user study. To fully explore the potential of force estimation, it is recommended to conduct more comprehensive data collections and user studies, involving a broader range of age groups, finger skin conditions, occupations, and medical conditions. By addressing these limitations, future research can further enhance the applicability and effectiveness of the proposed 3D relative force estimation method.

To synchronize image and force collection, a charging cable was connected to the phone. However, this setup may introduce variations in the capacitive images due to grounding effects and could slightly influence shear force measurements [14]. Additionally, participants completed the user study while holding the phone, a setup different from that used during the data collection phase.

While user performance in the study was satisfactory, the actual estimation accuracy was not evaluated under these conditions.

This study aims to propose a 3D relative force estimation method and compare it with other methods based on capacitive images. Consequently, estimation methods utilizing other sensors, such as barometers and accelerometers, are not evaluated in this paper. A key focus of future research will be to explore force estimation using multiple sensing techniques to enhance accuracy and robustness.

Changes in capacitive images can be influenced by both force and finger angle. Participants were instructed to maintain a constant finger angle during each task, but this may bring inconvenience to users in real-world settings. A potential solution to address this issue is to consider finger angle and force simultaneously.

The user study only employed the relative force between the first frame and subsequent frames in a captured sequence. However, as the sequence length increases, there is a higher likelihood of finger sliding. While a center alignment strategy was proposed to mitigate this problem, a more effective approach could involve updating the reference frame using a sliding window.

The study presented a simple force gesture recognition method based on thresholds. To further enhance recognition accuracy, more sophisticated algorithms like convolutional networks could be developed. Additionally, future research could explore real-world applications of force gestures.

We conducted the experiments exclusively using the Realme C11 (2021) smartphone due to the challenge of obtaining raw capacitive data without the cooperation of mobile phone manufacturers. However, one of the key objectives of our future research is to validate the effectiveness of our proposed method on a variety of mobile phones and tablets. By testing our method on different devices, we aim to ensure its applicability and robustness across a broader range of platforms.

9 Conclusion

In conclusion, this study presents a novel approach for estimating relative 3D forces applied by fingers on a touchscreen from two capacitive images, enabling continuous 3D force interactions. Empirical experiments demonstrate the superiority of the proposed relative force estimation method over SOTA absolute force estimation and indirect force estimation methods. The joint estimation of pressure and shear forces enhances accuracy by leveraging the inherent relationship between these two force components. The deployed network model, running on a commodity smartphone, showcases the practicality of the approach. Furthermore, a user study with four tasks confirms the effectiveness of the method in real-world scenarios, encompassing force gestures, pressure control, shear force control, and object manipulation using a combination of traditional touch input and force input. The study introduces a new set of force gestures and demonstrates that the proposed force estimation method achieves the highest recognition accuracy. The force control tasks illustrate the controllability and stability provided by the 3D relative force estimation method. Additionally, a 4DOF object manipulation task showcases the performance of the proposed method in a complex scenario, validating the capability of force input to be used in conjunction with conventional touch input.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 62376132 and 62321005.

References

- [1] Karan Ahuja, Paul Strelci, and Christian Holz. 2021. TouchPose: Hand Pose Prediction, Depth Estimation, and Touch Classification from Capacitive Images. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) (UIST '21). Association for Computing Machinery, New York, NY, USA, 997–1009. doi:10.1145/3472749.3474801
- [2] Ahmed Sabbir Arif and Wolfgang Stuerzlinger. 2013. Pseudo-Pressure Detection and Its Use in Predictive Text Entry on Touchscreens. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration* (Adelaide, Australia) (OzCHI '13). Association for Computing Machinery, New York, NY, USA, 383–392. doi:10.1145/2541016.2541024
- [3] Vincent Becker, Pietro Oldrati, Liliana Barrios, and Gábor Sörös. 2018. Touchsense: Classifying Finger Touches and Measuring their Force with an Electromyography Armband. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers* (Singapore, Singapore) (ISWC '18). Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3267242.3267250
- [4] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 1 (1995), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- [5] Tobias Boeckel, Sascha Sprott, Huy Viet Le, and Sven Mayer. 2019. Force Touch Detection on Capacitive Sensors Using Deep Neural Networks. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (Taipei, Taiwan) (MobileHCI '19). Association for Computing Machinery, New York, NY, USA, Article 42, 6 pages. doi:10.1145/3338286.3344389
- [6] Sebastian Boring, David Ledo, Xiang 'Anthony' Chen, Nicolai Marquardt, Anthony Tang, and Saul Greenberg. 2012. The Fat Thumb: Using the Thumb's Contact Size for Single-Handed Mobile Interaction. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services* (San Francisco, California, USA) (MobileHCI '12). Association for Computing Machinery, New York, NY, USA, 39–48. doi:10.1145/2371574.2371582
- [7] Davide Chicco. 2021. *Siamese Neural Networks: An Overview*. Springer US, New York, NY, 73–94. doi:10.1007/978-1-0716-0826-5_3
- [8] Frederic Choi, Sven Mayer, and Chris Harrison. 2021. 3D Hand Pose Estimation on Conventional Capacitive Touchscreens. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction* (Toulouse & Virtual, France) (MobileHCI '21). Association for Computing Machinery, New York, NY, USA, Article 3, 13 pages. doi:10.1145/3447526.3472045
- [9] Yongjie Duan, Ke He, Jianjiang Feng, Jiwen Lu, and Jie Zhou. 2022. Estimating 3D Finger Pose via 2D-3D Fingerprint Matching. In *27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) (IUI '22). Association for Computing Machinery, New York, NY, USA, 459–469. doi:10.1145/3490099.3511123
- [10] Yongjie Duan, Jinyang Yu, Jianjiang Feng, Ke He, Jiwen Lu, and Jie Zhou. 2023. 3D Finger Rotation Estimation from Fingerprint Images. *Proc. ACM Hum.-Comput. Interact.* 7, ISS, Article 431 (Nov 2023), 21 pages. doi:10.1145/3626467
- [11] Navid Fallahinia and Stephen A. Mascaró. 2022. Real-Time Tactile Grasp Force Sensing Using Fingernail Imaging via Deep Neural Networks. *IEEE Robotics and Automation Letters* 7, 3 (2022), 6558–6565. doi:10.1109/LRA.2022.3173751
- [12] Hyunjae Gil, DoYoung Lee, Seunggyu Im, and Ian Oakley. 2017. TriTap: Identifying Finger Touches on Smartwatches. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 3879–3890. doi:10.1145/3025453.3025561
- [13] Thomas R. Grieve, John M. Hollerbach, and Stephen A. Mascaró. 2015. 3-D Fingertip Touch Force Prediction Using Fingernail Imaging With Automated Calibration. *IEEE Transactions on Robotics* 31, 5 (2015), 1116–1129. doi:10.1109/TRO.2015.2459411
- [14] Tobias Grosse-Puppenthal, Christian Holz, Gabe Cohn, Raphael Wimmer, Oskar Bechtold, Steve Hodges, Matthew S. Reynolds, and Joshua R. Smith. 2017. Finding Common Ground: A Survey of Capacitive Sensing in Human-Computer Interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 3293–3315. doi:10.1145/3025453.3025808
- [15] Mark Hancock, Thomas ten Cate, and Sheelagh Carpendale. 2009. Sticky Tools: Full 6DOF Force-Based Interaction for Multi-Touch Tables. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Banff, Alberta, Canada) (ITS '09). Association for Computing Machinery, New York, NY, USA, 133–140. doi:10.1145/1731903.1731930
- [16] M.S. Hancock, F.D. Vernier, D. Wigdor, S. Carpendale, and Chia Shen. 2006. Rotation and Translation Mechanisms for Tabletop Interaction. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems* (TABLETOP

- '06). 8 pp.–. doi:10.1109/TABLETOP.2006.26
- [17] Chris Harrison and Scott Hudson. 2012. Using Shear as a Supplemental Two-Dimensional Input Channel for Rich Touchscreen Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (CHI '12). Association for Computing Machinery, New York, NY, USA, 3149–3152. doi:10.1145/2207676.2208730
- [18] Chris Harrison, Julia Schwarz, and Scott E. Hudson. 2011. TapSense: Enhancing Finger Interaction on Touch Surfaces. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) (UIST '11). Association for Computing Machinery, New York, NY, USA, 627–636. doi:10.1145/2047196.2047279
- [19] Ke He, Yongjie Duan, Jianjiang Feng, and Jie Zhou. 2022. Estimating 3D Finger Angle via Fingerprint Image. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 1, Article 14 (Mar 2022), 22 pages. doi:10.1145/3517243
- [20] Ke He, Chentao Li, Yongjie Duan, Jianjiang Feng, and Jie Zhou. 2024. TrackPose: Towards Stable and User Adaptive Finger Pose Estimation on Capacitive Touchscreens. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 161 (Jan 2024), 22 pages. doi:10.1145/3631459
- [21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [22] Seongkook Heo and Geehyuk Lee. 2011. Force Gestures: Augmenting touch Screen Gestures with Normal and Tangential Forces. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) (UIST '11). Association for Computing Machinery, New York, NY, USA, 621–626. doi:10.1145/2047196.2047278
- [23] Seongkook Heo and Geehyuk Lee. 2011. ForceTap: Extending the Input Vocabulary of Mobile Touch Screens by Adding Tap Gestures. In *Proceedings of the 13th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden) (MobileHCI '11). Association for Computing Machinery, New York, NY, USA, 113–122. doi:10.1145/2037373.2037393
- [24] Seongkook Heo and Geehyuk Lee. 2012. ForceDrag: Using Pressure as a Touch Input Modifier. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (Melbourne, Australia) (OzCHI '12). Association for Computing Machinery, New York, NY, USA, 204–207. doi:10.1145/2414536.2414572
- [25] Seongkook Heo and Geehyuk Lee. 2013. Indirect Shear Force Estimation for Multi-Point Shear Force Operations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI '13). Association for Computing Machinery, New York, NY, USA, 281–284. doi:10.1145/2470654.2470693
- [26] Christian Holz and Patrick Baudisch. 2011. Understanding Touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) (CHI '11). Association for Computing Machinery, New York, NY, USA, 2501–2510. doi:10.1145/1978942.1979308
- [27] Da-Yuan Huang, Ming-Chang Tsai, Ying-Chao Tung, Min-Lun Tsai, Yen-Ting Yeh, Liwei Chan, Yi-Ping Hung, and Mike Y. Chen. 2014. TouchSense: Expanding Touchscreen Input Vocabulary Using Different Areas of Users' Finger Pads. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI '14). Association for Computing Machinery, New York, NY, USA, 189–192. doi:10.1145/2556288.2557258
- [28] Mengting Huang, Kazuyuki Fujita, Kazuki Takashima, Taichi Tsuchida, Hiroyuki Manabe, and Yoshifumi Kitamura. 2019. ShearSheet: Low-Cost Shear Force Input with Elastic Feedback for Augmenting Touch Interaction. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) (ISS '19). Association for Computing Machinery, New York, NY, USA, 77–87. doi:10.1145/3343055.3359717
- [29] Zeyuan Huang, Cangjun Gao, Haiyan Wang, Xiaoming Deng, Yu-Kun Lai, Cuixia Ma, Sheng-feng Qin, Yong-Jin Liu, and Hongan Wang. 2024. SpecFingers: Finger Identification and Error Correction on Capacitive Touchscreens. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 8 (Mar 2024), 28 pages. doi:10.1145/3643559
- [30] Sungjae Hwang, Andrea Bianchi, and Kwang-yun Wahn. 2013. VibPress: Estimating Pressure Input using Vibration Absorption on Mobile Devices. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Munich, Germany) (MobileHCI '13). Association for Computing Machinery, New York, NY, USA, 31–34. doi:10.1145/2493190.2493193
- [31] Kaori Ikematsu and Shota Yamanaka. 2020. ScraTouch: Extending Interaction Technique Using Fingernail on Unmodified Capacitive Touch Surfaces. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 81 (Sep 2020), 19 pages. doi:10.1145/3411831
- [32] Y. Kurita, A. Ikeda, Jun Ueda, and T. Ogasawara. 2005. A Fingerprint Pointing Device Utilizing the Deformation of the Fingertip During the Incipient Slip. *IEEE Transactions on Robotics* 21, 5 (2005), 801–811. doi:10.1109/TRO.2005.847568
- [33] Huy Viet Le, Thomas Kosch, Patrick Bader, Sven Mayer, and Niels Henze. 2018. PalmTouch: Using the Palm as an Additional Input Modality on Commodity Smartphones. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3173934
- [34] Huy Viet Le, Sven Mayer, and Niels Henze. 2019. Investigating the Feasibility of Finger Identification on Capacitive Touchscreens using Deep Learning. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Rey, California) (IUI '19). Association for Computing Machinery, New York, NY, USA, 637–649. doi:10.1145/3301275.3302295
- [35] Borham Lee, Hyunjeong Lee, Soo-Chul Lim, Hyungkeew Lee, Seungju Han, and Joonah Park. 2012. Evaluation of Human Tangential Force Input Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (CHI '12). Association for Computing Machinery, New York, NY, USA, 3121–3130. doi:10.1145/2207676.2208727
- [36] Jingbo Liu, Oscar Kin-Chung Au, Hongbo Fu, and Chiew-Lan Tai. 2012. Two-Finger Gestures for 6DOF Manipulation of 3D Objects. *Computer Graphics Forum* 31, 7 (2012), 2047–2055. doi:10.1111/j.1467-8659.2012.03197.x
- [37] Anthony Martinet, Gery Casiez, and Laurent Grisoni. 2012. Integrality and Separability of Multitouch Interaction Techniques in 3D Manipulation Tasks. *IEEE Transactions on Visualization and Computer Graphics* 18, 3 (2012), 369–380. doi:10.1109/TVCG.2011.129
- [38] Sven Mayer, Perihan Gad, Katrin Wolf, Pawel W. Woźniak, and Niels Henze. 2017. Understanding the Ergonomic Constraints in Designing for Touch Surfaces. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) (MobileHCI '17). Association for Computing Machinery, New York, NY, USA, Article 33, 9 pages. doi:10.1145/3098279.3098537
- [39] Sven Mayer, Huy Viet Le, and Niels Henze. 2017. Estimating the Finger Orientation on Capacitive Touchscreens Using Convolutional Neural Networks. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (Brighton, United Kingdom) (ISS '17). Association for Computing Machinery, New York, NY, USA, 220–229. doi:10.1145/3132272.3134130
- [40] Yuriko Nakai, Shinya Kudo, Ryuta Okazaki, Hiroyuki Kajimoto, and Hidenori Kuribayashi. 2014. Detection of Tangential Force for a Touch Panel Using Shear Deformation of the Gel. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI EA '14). Association for Computing Machinery, New York, NY, USA, 2353–2358. doi:10.1145/2559206.2581149
- [41] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. 2015. Sensing Touch Force using Active Acoustic Sensing. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction* (Stanford, California, USA) (TEI '15). Association for Computing Machinery, New York, NY, USA, 355–358. doi:10.1145/2677199.2680585
- [42] Philip Quinn. 2019. Estimating Touch Force with Barometric Pressure Sensors. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3290605.3300919
- [43] Philip Quinn, Wenxin Feng, and Shumin Zhai. 2021. *Deep Touch: Sensing Press Gestures from Touch Image Sequences*. Springer International Publishing, Cham, 169–192. doi:10.1007/978-3-030-82681-9_6
- [44] Anne Roudaut, Eric Lecolinet, and Yves Guiard. 2009. MicroRolls: Expanding Touch-Screen Input Vocabulary by Distinguishing Rolls vs. Slides of the Thumb. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) (CHI '09). Association for Computing Machinery, New York, NY, USA, 927–936. doi:10.1145/1518701.1518843
- [45] Robin Schweigert, Jan Leusmann, Simon Hagenmayer, Maximilian Weiß, Huy Viet Le, Sven Mayer, and Andreas Bulling. 2019. KnuckleTouch: Enabling Knuckle Gestures on Capacitive Touchscreens using Deep Learning. In *Proceedings of Mensch Und Computer 2019* (Hamburg, Germany) (MuC '19). Association for Computing Machinery, New York, NY, USA, 387–397. doi:10.1145/3340764.3340767
- [46] Paul Strelciak and Christian Holz. 2021. CapContact: Super-resolution Contact Areas from Capacitive Touchscreens. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 289, 14 pages. doi:10.1145/3411764.3445621
- [47] Ryosuke Takada, Wei Lin, Toshiyuki Ando, Buntarou Shizuki, and Shin Takahashi. 2017. A Technique for Touch Force Sensing using a Waterproof Device's Built-in Barometer. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI EA '17). Association for Computing Machinery, New York, NY, USA, 2140–2146. doi:10.1145/3027063.3053130
- [48] Jamie Ullerich, Maximiliane Windl, Andreas Bulling, and Sven Mayer. 2023. ThumbPitch: Enriching Thumb Interaction on Mobile Touchscreens using Deep Learning. In *Proceedings of the 34th Australian Conference on Human-Computer Interaction* (Canberra, ACT, Australia) (OzCHI '22). Association for Computing Machinery, New York, NY, USA, 58–66. doi:10.1145/3572921.3572925
- [49] Yoichi Watanabe, Yasutoshi Makino, Katsunari Sato, and Takashi Maeno. 2012. Contact Force and Finger Angles Estimation for Touch Panel by Detecting Transmitted Light on Fingernail. In *Haptics: Perception, Devices, Mobility, and Communication*, Poika Isokoski and Jukka Springare (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 601–612.

- [50] Frank Wilcoxon. 1992. *Individual Comparisons by Ranking Methods*. Springer New York, New York, NY, 196–202. doi:10.1007/978-1-4612-4380-9_16
- [51] Robert Xiao, Gierad Laput, and Chris Harrison. 2014. Expanding the Input Expressivity of Smartwatches with Mechanical Pan, Twist, Tilt and Click. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (*CHI '14*). Association for Computing Machinery, New York, NY, USA, 193–196. doi:10.1145/2556288.2557017
- [52] Robert Xiao, Julia Schwarz, and Chris Harrison. 2015. Estimating 3D Finger Angle on Commodity Touchscreens. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces* (Madeira, Portugal) (*ITS '15*). Association for Computing Machinery, New York, NY, USA, 47–50. doi:10.1145/2817721.2817737
- [53] Jinyang Yu, Jianjiang Feng, and Jie Zhou. 2023. PrintShear: Shear Input Based on Fingerprint Deformation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 2, Article 81 (Jun 2023), 22 pages. doi:10.1145/3596257
- [54] Vadim Zaliva. 2012. 3D Finger Posture Detection and Gesture Recognition on Touch Surfaces. In *2012 12th International Conference on Control Automation Robotics & Vision (ICARCV)*. 359–364. doi:10.1109/ICARCV.2012.6485185