TeachingBot: Robot Teacher for Human Handwriting

Zhimin Hou[1]*, Cunjun Yu[2]*, David Hsu[2,3], Haoyong Yu[1]

***Abstract*—Teaching and learning physical skills often require one-on-one interaction, making it difficult to scale up, as there are not enough human teachers. Robots offer an attractive alternative. This paper presents *TeachingBot*, an adaptive robotic system that teaches handwriting to human learners through physical interaction. Robot teaching poses two major challenges: (i) adapting to the individual handwriting style of the learner and (ii) maintaining an engaging learning experience. For the first challenge, TeachingBot uses a probabilistic model to capture the learner's writing style from their writing samples. Drawing on the insight that effective teaching balances standardization with individuality, the system generates a personalized teaching trajectory that aligns with the learner's natural writing. For the second challenge, TeachingBot employs variable impedance control to guide the learner, dynamically adjusting the strength of physical guidance based on the learner's performance. Human-subject experiments with 15 participants demonstrate the effectiveness of TeachingBot, showing clear improvement in learners' handwriting and engagement over baseline methods.**

I. Introduction

Robots play a crucial role in various physical interaction tasks with humans [1]. While robots have traditionally served as learners [2, 3], collaborators [4], or assistants [5], they hold great potential as teachers for humans [6]. *How can robots provide effective, personalized physical guidance for teaching physical skills when human teachers are scarce?* Specifically, this work focuses on handwriting, an essential physical skill in daily life. The robot serves as the *teacher*, providing guidance through physical interaction with humans to teach handwriting.

Teaching humans to write effectively presents two key challenges. First, human learners exhibit diverse writing styles for the same character [7, 8, 9], with variations in stroke length, position, and direction—especially evident in Chinese characters [8]. Learners naturally prefer guidance that accommodates their individual writing styles. Second, determining the appropriate level of guidance is difficult: insufficient support fails to facilitate learning, while excessive assistance can create dependency [5] and hinder skill acquisition [6]. A successful robot teaching system must therefore (1) adapt to individual writing styles and (2) provide an appropriate level of physical guidance that promotes active engagement.

We present ***TeachingBot***, an adaptive robot teaching system designed to teach Chinese character writing through physical guidance, as illustrated in Fig. 1. Our key insight is that effective teaching should balance standardization with individuality since guidance is more effective when grounded in the human learner's own writing style. TeachingBot achieves this through a two-step adaptive approach.

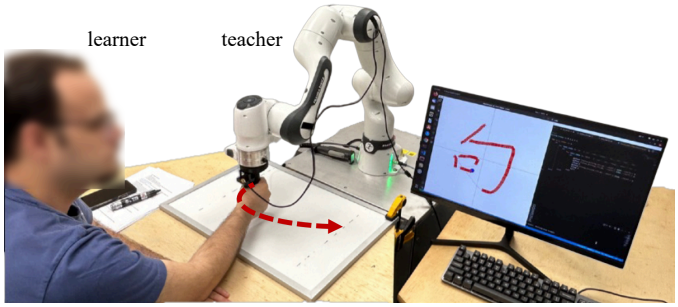


Fig. 1: *TeachingBot*: the robot teacher guides the learner in writing Chinese characters, adapting to the learner's preferences and proficiency.

First, the system captures the human learner's writing style by collecting their writing trajectories and modeling them probabilistically. Drawing inspiration from learning-from-demonstration methods [10, 11, 12], we use a mixture of Gaussian distributions to encode the human learner's writing style, capturing both the average writing trajectory and its variations. By combining this model with the reference character the robot aim to teach, we generate a teaching trajectory that respects the human learner's writing preferences.

Second, the robot executes the teaching trajectory through compliant physical interaction using variable impedance control [13, 14, 15]. The human learner holds the robotic arm's end effector, which guides them through the generated teaching trajectory (see Fig. 1). The system dynamically adjusts the guidance force: applying stronger correction when the human learner strays far from the reference, and reducing assistance when deviations are minor to encourage exploration and maintain engagement [5, 16, 17]. Through this adaptive strategy, TeachingBot creates a personalized, engaging learning experience that respects individual writing styles while guiding learners toward reference character style. This approach is especially suitable in scenarios where detailed models of human learning dynamics are unavailable [18, 6].

It is crucial to distinguish between *robot learning* and *robot teaching*. While robot learning focuses on robots acquiring new capabilities, robot teaching, as exemplified by TeachingBot, positions robots as instructors for human learners. Through human-subject experiments with 15 participants with varying levels of Chinese writing proficiency, we demonstrate

* Equal contribution.

The authors are associated with the [1]Department of Biomedical Engineering, [2]School of Computing, [3]Smart System Institute, National University of Singapore.

In Section IV, all subjects gave informed consent and testing was approved by National University of Singapore Institutional Review Board (NUS-IRB-2023-875).

IRB has been acquired

that TeachingBot significantly improves handwriting quality compared to two baseline methods. Participants showed greater accuracy in replicating both the overall structure and fine-grained stroke details of characters. Additionally, increased interaction forces during training indicate enhanced learner engagement, highlighting the effectiveness of adaptive physical guidance in promoting active skill acquisition.

One limitation of this work is that it does not explicitly address long-term retention or how well the acquired skills transfer beyond the immediate training sessions.

II. Related Work

A. Teaching Algorithms

Various algorithms have been used for human learning, with successes in areas such as crowd classification and concept acquisition [19, 20, 21, 22]. However, mastering intricate motor control skills through mere visual or verbal cues remains a challenge. Recent advances integrate skill discovery techniques from reinforcement learning to establish curricula based on skill decomposition, enhancing human motor skill acquisition [18]. The emergence of advanced language models has introduced language correction as a tool to facilitate human learning [23]. The concept of robot teaching has emerged in the realm of robot-assisted learning [24, 7]. Robots can physically facilitate human learning by using specified dynamics [6, 25, 26]. In this work, instead of only providing physical guidance, we aim to achieve adaptive teaching by learning and incorporating the style of the learner in the physical guide.

B. Robot-assisted Training

While social robots have successfully facilitated language learning and writing skills [27], physical human-robot interaction (pHRI) proves more effective for tasks requiring subtle motor control by enabling direct physical guidance [15]. Robot-assisted training has been extensively studied in various rehabilitation applications [13]. The key challenges in pHRI-based learning systems lie in trajectory generation and interaction control. For trajectory generation, probabilistic methods like Gaussian Mixture Models (GMM)/Regression (GMR) and Dynamic Movement Primitives (DMP) [10, 11, 28] enable robots to generate smooth, human-like reference trajectories. Advanced approaches such as Probabilistic DMP (ProMP) and Gaussian Process (GP) models [29, 12, 3] further improve adaptability through basis function representations. While these methods have proven effective in robot-assisted rehabilitation [13, 30], they primarily focus on robots learning from experts/therapists rather than actively optimizing trajectory generation for human skill acquisition.

The effectiveness of physical guidance also depends critically on the adjustment of the interaction intensity [13]. Research shows that active user engagement enhances neural plasticity and learning efficiency [31, 5]. Variable impedance controllers have emerged as key tools for modulating this interaction [32, 13], with approaches ranging from error-based parameter adjustment [5, 14] to model-free learning-based methods [16, 17, 33]. However, these approaches have only shown effectiveness in promoting active participation, while failing to simultaneously accommodate individual learning preferences and encourage user participation during interaction control. TeachingBot addresses these challenges by dynamically adjusting the teaching intensity based on both the learner's style and the real-time writing performance.

III. Methodology

A. Problem Formulation

We build on prior work [6, 18] that models the *target task* (the skill to be taught) as a Markov Decision Process (MDP) and the *teaching task* (the teaching process) as a Partially Observable MDP (POMDP). Here, the target task is writing Chinese characters, and the teaching task is to guide humans in learning this skill. The *teaching policy* influences learner performance through interactive actions, such as generating reference trajectories and adjusting physical interaction impedance. The objective of TeachingBot is to adaptively teach learners to write the reference character. Defining the POMDP elements is challenging due to the lack of an accurate human learning model, so we adopt a simplified solution rather than solving the full POMDP. TeachingBot adjusts its support to the human learner's needs purely based on the current observation of the human learner's writing style encoded by the probabilistic learning method GMM/GMR and writing performance.

Building upon the formulation for pHRI [5], the dynamics of the *robot teacher* are expressed in the end-effector space as

$$\boldsymbol{M}(\boldsymbol{x})\ddot{\boldsymbol{x}} + \boldsymbol{C}(\boldsymbol{x}, \dot{\boldsymbol{x}})\dot{\boldsymbol{x}} + \boldsymbol{G}(\boldsymbol{x}) = \boldsymbol{F}_h + \boldsymbol{F}_r \tag{1}$$

where $\boldsymbol{x}$ and $\dot{\boldsymbol{x}}$ are the actual position and velocity, respectively. $\boldsymbol{F}_r$ is the assistive force provided by *the robot teacher*. $\boldsymbol{F}_h$ is the *human learner* actively applied force, typically measured to evaluate the engagement in pHRI tasks [5, 15, 17]. A variable impedance controller $\pi(\boldsymbol{u}|\boldsymbol{x}, \dot{\boldsymbol{x}}; \boldsymbol{x}_d, \dot{\boldsymbol{x}}_d, \boldsymbol{\mathcal{K}}_d, \boldsymbol{\mathcal{B}}_d)$ is implemented to provide corrective guidance given the reference trajectory and impedance parameters [15]. $\boldsymbol{x}_d \in \mathbb{R}^3$ and $\dot{\boldsymbol{x}}_d \in \mathbb{R}^3$ are the reference position and velocity, respectively. At each actuation step with a sample interval T_s, the control input $\boldsymbol{u}$ is designed as

$$\boldsymbol{u} = \boldsymbol{J}^T\big[-\boldsymbol{\mathcal{K}}_d(\boldsymbol{x} - \boldsymbol{x}_d) - \boldsymbol{\mathcal{B}}_d(\dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_d) + \boldsymbol{F}_{fd}\big], \tag{2}$$

where $\boldsymbol{J}$ is the Jacobian matrix. $\boldsymbol{F}_{fd}$ is introduced to compensate for the robot dynamics as formulated in (1). $\boldsymbol{\mathcal{K}}_d \in \mathbb{R}^{3\times3}$ and $\boldsymbol{\mathcal{B}}_d \in \mathbb{R}^{3\times3}$ are the reference stiffness and damping. They are regulated to adjust the assistance level provided by the *robot teacher* and the engagement level of the *human learner* [13, 14, 17].

B. Overview of TeachingBot

The overview of TeachingBot is depicted in Fig. 2. The reference character is denoted by the variable $\boldsymbol{c} \in \boldsymbol{\mathcal{C}}$. $\boldsymbol{s}_i \in \boldsymbol{\mathcal{S}}_{\boldsymbol{c}}$ represents i-th stroke of character $\boldsymbol{c}$. The dataset of reference Chinese characters, $\boldsymbol{\mathcal{C}}$, includes images of each character $\{\boldsymbol{I}_c\}$ for the human users to learn (see Fig. 2(a)). $\boldsymbol{\mathcal{S}}_{\boldsymbol{c}}$ includes all stroke images of the character $\boldsymbol{c}$. To facilitate the generation of robot teaching trajectories, reference waypoints $\{\boldsymbol{\chi}_{\boldsymbol{c}}^n\}_{n=1}^N$ of all strokes are extracted from the image of the reference character based on image processing method, N is the number of waypoints and $\boldsymbol{\chi}_c^n$ denotes the n-th reference waypoint. As

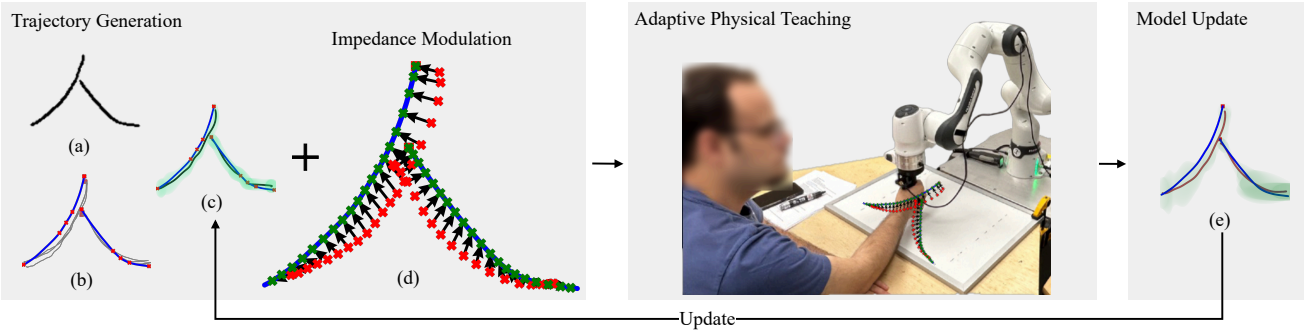


Fig. 2: Overview of TeachingBot. First, a reference character (a) is selected, and waypoints (b) are extracted from both the reference and the learner's handwriting. Their differences guide initial impedance modulation (d), where the blue/green path shows the reference trajectory/waypoints, the red cross indicates the learner's mean writing waypoints, and black arrows represent the potential correction forces used to modulate the impedance parameters for robot control. The robot then physically guides and teaches the learner. After teaching, a GMM-GMR (e) updates the learner's writing style, and GMR-GP generates a new training trajectory for the next teaching iteration (black in c).

Algorithm 1 TeachingBot

Input: Character set $\boldsymbol{\mathcal{C}}$, number of teaching iterations M, number of evaluation iterations L
Initialize: $\boldsymbol{\mathcal{D}}_V^0 \leftarrow \emptyset$, $\boldsymbol{\mathcal{D}}_L^0 \leftarrow \emptyset$, $\boldsymbol{\mathcal{K}}_s^0 \leftarrow \mathbf{0}$, $\boldsymbol{\mathcal{B}}_s^0 \leftarrow \mathbf{0}$, GMR-GP
for each character $\boldsymbol{c} \in \boldsymbol{\mathcal{C}}$ **do**
  Extract reference waypoints $\boldsymbol{\chi}_c$ from image $\boldsymbol{I}_c$
  /// Pre-test Phase ///
  for $l = 1$ to L **do**
    Collect learner's trial image $\boldsymbol{I}_a^l$ and extract waypoints $\boldsymbol{\chi}_a^l$
  end for
  Encode learner's initial writing style $\overline{\boldsymbol{\chi}}_a^I$ from $\{\boldsymbol{\chi}_a^l\}_{l=1}^L$
  Initialize robot impedance parameters $\boldsymbol{\mathcal{K}}_r$, $\boldsymbol{\mathcal{B}}_r$
  /// Teaching Phase ///
  for $m = 1$ to M **do**
    for each stroke $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}_{\boldsymbol{c}}$ **do**
      Update learner's writing style from $\boldsymbol{\mathcal{D}}_L^{m-1}$
      Retrieve H training via-points from $\boldsymbol{\mathcal{D}}_V^m$
      Generate teaching trajectory $\boldsymbol{\chi}_d^m$ via GMR-GP model
      Compute impedance: stiffness $\boldsymbol{\mathcal{K}}_d^m$, damping $\boldsymbol{\mathcal{B}}_d^m$
      Apply physical guidance using VIC (Eq. 2)
      Record learner's actual writing trajectory $\boldsymbol{\chi}_a^m$
      Update dataset: $\boldsymbol{\mathcal{D}}_L^m \leftarrow \boldsymbol{\mathcal{D}}_L^{m-1} \cup \{\boldsymbol{t}, \boldsymbol{\chi}_a^m\}$, $\boldsymbol{\mathcal{D}}_V^m \leftarrow \emptyset$
    end for
  end for
  /// Evaluation Phase ///
  for $l = 1$ to L **do**
    Collect evaluation image $\boldsymbol{I}_a^l$ and extract waypoints $\boldsymbol{\chi}_a^l$
  end for
  Encode learner's writing style $\overline{\boldsymbol{\chi}}_a^E$ from $\{\boldsymbol{\chi}_a^l\}_{l=1}^L$
end for

illustrated in Fig. 2(b), the reference waypoints of two strokes are extracted and plotted as the blue lines. Given a reference character $\boldsymbol{c}$, our TeachingBot is implemented following three phases, as presented in Algorithm 1. During the pre-test and evaluation phases, the human learner is instructed to write the reference character for L iterations without the physical guidance of the robot. During the teaching phase, a complete round of writing the given character is defined as one *teaching iteration*. Robot teaching is repeated for M iterations, at m-th teaching iteration ($m \in M$), the previous L actual writing trajectories are collected as $\{\boldsymbol{\mathcal{T}}_a^l\}_{l=1}^L$, which are downsampled into actual writing waypoints $\{\{\boldsymbol{\chi}_a^{l,n}\}_{n=1}^N\}_{l=1}^L$. A dataset comprising all time-driven writing waypoints, $\mathcal{D}_L^m = \{\{\boldsymbol{t}_n^l,\ \boldsymbol{\chi}_a^{l,n}\}_{n=1}^N\}_{l=1}^L$, is then constructed to encode the human learner's writing style. Afterwards, the key via-points are extracted from the reference waypoints and stored in $\mathcal{D}_V^m$, defined as training via-points. A GMR-GP model, $f(\boldsymbol{\chi}_r|\boldsymbol{t}, \mathcal{D}_L^m, \mathcal{D}_V^m)$, is trained to generate training waypoints $\{\boldsymbol{\chi}_d^{m,n}\}_{n=1}^N$ based on the encoded human learner's writing style and the training via-points. Finally, the training waypoints are interpolated to create the reference trajectory integrating with the adapted impedance parameters for physical teaching interactive control (see Section III-F). As illustrated in Fig. 2, the human learner would grasp the robot's handle, and the robot teacher guides the human learner to write the reference character by following a generated reference trajectory. Afterwards, the reference trajectory and impedance parameters will be updated for the robot teacher to personalize the teaching task.

C. Human Learner Writing Style Learning

For the m-th teaching iteration, the human learner's writing style is encoded using GMM/GMR with Z components. The joint distribution of the input time $\boldsymbol{\xi}_i = \boldsymbol{t} \in \mathbb{R}^{d_i}$, and the output writing waypoint $\boldsymbol{\xi}_o = \boldsymbol{\chi} \in \mathbb{R}^{d_o}$ is modelled as

$$\mathcal{P}(\boldsymbol{\xi}) = \sum\nolimits_{z=1}^{Z} h_z \mathcal{N}(\boldsymbol{\xi}; \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z), \boldsymbol{\xi} = [\boldsymbol{\xi}_i, \boldsymbol{\xi}_o]^T, \tag{3}$$

where h_z, $\boldsymbol{\mu}_z$, and $\boldsymbol{\Sigma}_z$ are the prior probability, mean, and covariance of the z-th Gaussian component, respectively. These parameters are optimized using the Expectation-Maximization algorithm given the dataset $\mathcal{D}_L^{m-1}$[11, 12]. Therefore, the writing style of human learner is represented by a probabilistic writing waypoints $\{\widehat{\boldsymbol{\chi}}_a^n\}_{n=1}^N$. Each waypoint $\widehat{\boldsymbol{\chi}}_a^n$ is retrieved from a conditional Gaussian distribution $\mathcal{P}(\widehat{\boldsymbol{\chi}}_a^n|\boldsymbol{t}_n) = \mathcal{N}(\widehat{\boldsymbol{\mu}}(\boldsymbol{t}_n), \widehat{\boldsymbol{\Sigma}}(\boldsymbol{t}_n))$. $\widehat{\boldsymbol{\mu}}(\boldsymbol{t}_n)$ and $\widehat{\boldsymbol{\Sigma}}(\boldsymbol{t}_n)$ are the conditional mean and covariance, respectively. The covariance $\widehat{\boldsymbol{\Sigma}}(\boldsymbol{t}_n)$ encapsulates the variability of the human learner's potential writing trajectories, providing a measure of uncertainty and flexibility in their writing style.

As depicted in Fig. 2(a) and Fig. 2(b), L=3 writing waypoints of the reference character $\boldsymbol{c}$ are represented by gray lines. The GMMs, with Z=8 components, are optimized and visualized as the green ellipses in Fig. 2(e). The mean writing waypoints and variance, derived from GMMs, are depicted as the green line and green shaded region in Fig. 2(f).

D. Training Via-points Extraction

The curvature-based trajectory compression is employed to extract the training via-points from the reference waypoints $\boldsymbol{\chi}_c$. Particularly, we retain waypoints where the curvature (change in direction) is highest, as these waypoints are likely

to represent features of the reference character. Given the reference waypoints $\{\boldsymbol{\chi}_{\boldsymbol{c}}^n\}_{n=1}^N$, calculate curvature κ_n for each interior waypoint $\boldsymbol{\chi}_{\boldsymbol{c}}^n$ using the first order derivative $\dot{\boldsymbol{\chi}}_{\boldsymbol{c}}^n$ and second order derivative $\ddot{\boldsymbol{\chi}}_{\boldsymbol{c}}^n$. The curvature can be computed as $\kappa_n = \ddot{\boldsymbol{\chi}}_{\boldsymbol{c}}^n/(1+(\dot{\boldsymbol{\chi}}_{\boldsymbol{c}}^n)^2)^{1.5}$. Then, H waypoints with the highest curvature values are selected as training via-points, indicating significant changes in trajectory direction. As introduced in Algorithm 1, during the m-th teaching iteration, the training via-points are obtained and stored in $\mathcal{D}_V^m = \{\boldsymbol{t}_h^m, \boldsymbol{\chi}_c^h\}_{h=1}^H$ ($H << N$). As depicted in Fig. 2(b), the extracted training via-points ($H=5$) are visualized as red scatter points.

E. Training Waypoints Generation

For the m-th teaching iteration, given the previous writing waypoints in $\mathcal{D}_L^{m-1}$, the training waypoints need to be generated to maintain the current writing style and follow the instruction of training waypoints in $\mathcal{D}_V^m$. Multi-output GP (MOGP) is employed to fit the deterministic relationship $\boldsymbol{\xi}_o = f(\boldsymbol{\xi}_i) + \boldsymbol{\epsilon}_t$ from the input $\boldsymbol{\xi}_i = \boldsymbol{t}$ to vector-valued output $\boldsymbol{\xi}_o = \boldsymbol{\chi}$, $\boldsymbol{\epsilon}_t = [\boldsymbol{\epsilon}_t^1, \boldsymbol{\epsilon}_t^p, \cdots, \boldsymbol{\epsilon}_t^{d_o}]$, $\boldsymbol{\epsilon}_t^p \sim \mathcal{N}(0, \sigma_p^2)$. The distribution of $\boldsymbol{\chi}$ input $\boldsymbol{t}$ is given by $\boldsymbol{\chi}(\boldsymbol{t}) \sim \mathcal{GP}(\boldsymbol{\mu}(\boldsymbol{t}), \boldsymbol{k}(\boldsymbol{t}, \boldsymbol{t}'))$, where $\boldsymbol{\mu}(\cdot): \mathbb{R}^{d_i} \to \mathbb{R}^{d_o}$ and $\boldsymbol{k}(\cdot,\cdot): \mathbb{R}^{d_i} \times \mathbb{R}^{d_i} \to \mathbb{R}^{d_o} \times \mathbb{R}^{d_o}$ are the mean and kernel function. The joint distribution of observed samples and the predicted output $\boldsymbol{\chi}^*$ of the input $\boldsymbol{t}_*$ is given as follows

$$\begin{bmatrix} \boldsymbol{\chi}^{1:N} \\ \boldsymbol{\chi}^* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}(\boldsymbol{t}_{1:N}) \\ \boldsymbol{\mu}(\boldsymbol{t}_*) \end{bmatrix}, \begin{bmatrix} \boldsymbol{K}(\boldsymbol{t},\boldsymbol{t}) + \boldsymbol{\Sigma}_\epsilon & \boldsymbol{K}(\boldsymbol{t},\boldsymbol{t}_*) \\ \boldsymbol{K}(\boldsymbol{t}_*,\boldsymbol{t}) & \boldsymbol{K}(\boldsymbol{t}_*,\boldsymbol{t}_*) \end{bmatrix} \right), \tag{4}$$

where $\boldsymbol{\chi}^{1:N}$ are the observed values correspond to the input $\boldsymbol{t}_{1:N}$. $\boldsymbol{K}(\boldsymbol{t},\boldsymbol{t}) \in \mathbb{R}^{Nd_o \times Nd_o}$, $\boldsymbol{K}(\boldsymbol{t},\boldsymbol{t}_*) \in \mathbb{R}^{Nd_o \times d_o}$, $\boldsymbol{K}(\boldsymbol{t}_*,\boldsymbol{t}) \in \mathbb{R}^{d_o \times Nd_o}$, and $\boldsymbol{K}(\boldsymbol{t}_*,\boldsymbol{t}_*) \in \mathbb{R}^{d_o \times d_o}$ are the Gram matrices that all elements are calculated using the kernel function over all input pairs $(\boldsymbol{t},\boldsymbol{t})$. Similarly to [12, 2], the kernel function $\boldsymbol{k}(\boldsymbol{t},\boldsymbol{t}') = \sum_{q=1}^Q \boldsymbol{\Xi}_q \boldsymbol{k}_q(\boldsymbol{t},\boldsymbol{t}')$ is designed based on the Linear Model of Coregionalization (LMC) assumption. $\boldsymbol{\Xi}_q \in \mathbb{R}^{d_o \times d_o}$ is a positive semi-definite coregionalization matrix. $\boldsymbol{k}_q(\boldsymbol{t},\boldsymbol{t}')$ is the scalar kernel function.

The posterior distribution of the human learner's writing waypoints is derived as a multivariate Gaussian distribution (MGD), as $\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \mathcal{D}_L^{m-1}) \sim \mathcal{N}(\boldsymbol{\mu}_L^*, \boldsymbol{\Sigma}_L^*)$. The mean and covariance $\boldsymbol{\mu}_L^*$ and $\boldsymbol{\Sigma}_L^*$ are computed as follows

$$\begin{aligned} \boldsymbol{\mu}_L^* &= \boldsymbol{\mu}(\boldsymbol{t}_*) + \boldsymbol{K}(\boldsymbol{t}_*,\boldsymbol{t})(\boldsymbol{K}(\boldsymbol{t},\boldsymbol{t}) + \boldsymbol{\Sigma}_\epsilon)^{-1}(\boldsymbol{\chi}^{1:N} - \boldsymbol{\mu}(\boldsymbol{t})) \\ \boldsymbol{\Sigma}_L^* &= \boldsymbol{K}(\boldsymbol{t}_*,\boldsymbol{t}_*) - \boldsymbol{K}(\boldsymbol{t}_*,\boldsymbol{t})(\boldsymbol{K}(\boldsymbol{t},\boldsymbol{t}) + \boldsymbol{\Sigma}_\epsilon)^{-1}\boldsymbol{K}(\boldsymbol{t},\boldsymbol{t}_*), \end{aligned} \tag{5}$$

where the training via-points in $\mathcal{D}_V^m$ are treated as new observations to complement the previous observations in $\mathcal{D}_L^{m-1}$. Unlike standard MOGP, GMR-GP model replaces the prior mean value $\boldsymbol{\mu}(\boldsymbol{t}_*)$ with the estimated GMR mean value. The kernel function $\boldsymbol{k}(\boldsymbol{t},\boldsymbol{t}') = \sum_{z=1}^Z h_z(\boldsymbol{t}) h_z(\boldsymbol{t}') \widehat{\boldsymbol{\Sigma}}_z \boldsymbol{k}_z(\boldsymbol{t},\boldsymbol{t}')$ is designed based on the learned variability. $h_z(\boldsymbol{t})$ and $\widehat{\boldsymbol{\Sigma}}_z$ are the derived responsibilities and componentwise conditional covariance matrix calculated by (3). The MGD for the training waypoints generation is rewritten as follows

$$\begin{aligned} &\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \mathcal{D}_L^{m-1}, \mathcal{D}_V^m) \propto \mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \mathcal{D}_L^{m-1})\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \mathcal{D}_V^m), \\ &\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \mathcal{D}_L^{m-1}, \mathcal{D}_V^m) \sim \mathcal{N}(\boldsymbol{\mu}_{L-V}^m, \boldsymbol{\Sigma}_{L-V}^m), \end{aligned} \tag{6}$$

where $\boldsymbol{\mu}_{L-V}^m$ and $\boldsymbol{\Sigma}_{L-V}^m$ are derived based on $\{\boldsymbol{\mu}_L^*, \boldsymbol{\Sigma}_L^*, \boldsymbol{\mu}_V^*, \boldsymbol{\Sigma}_V^*\}$ [2]. $\boldsymbol{\mu}_L^*$ and $\boldsymbol{\Sigma}_L^*$ are the mean and covariance of $\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \mathcal{D}_L^{m-1})$; $\boldsymbol{\mu}_V^*$ and $\boldsymbol{\Sigma}_V^*$ are the mean and covariance of $\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \mathcal{D}_V^m)$. Subsequently, the training waypoints $\{\boldsymbol{\chi}_d^{m,n}\}_{n=1}^N$ are sampled from the learned posterior distribution in (6) given $\{\boldsymbol{t}_n\}_{n=1}^N$. Using the training via-points shown in Fig. 2(b) and encoded writing styles in Fig. 2(e), the sampled training waypoints are illustrated by the black line in Fig. 2(c).

F. Physical Teaching Interactive Control

The actual writing trajectory and interaction force trajectory are respectively collected as $\boldsymbol{\mathcal{T}}_a = [\boldsymbol{x}(0), \boldsymbol{x}(T_s), \ldots, \boldsymbol{x}(\Delta \boldsymbol{t}_c)]$ and $\boldsymbol{\Gamma}_a = [\boldsymbol{F}_h(0), \boldsymbol{F}_h(T_s), \ldots, \boldsymbol{F}_h(\Delta \boldsymbol{t}_c)]$. $\Delta \boldsymbol{t}_c$ is writing duration for the given reference character $\boldsymbol{c}$. After obtaining the training waypoints $\{\boldsymbol{\chi}_d^{m,n}\}_{n=1}^N$ from the m-th teaching iteration, the reference trajectory $\boldsymbol{\mathcal{T}}_d = [\boldsymbol{x}_d(0), \boldsymbol{x}_d(T_s), \cdots, \boldsymbol{x}_d(\Delta \boldsymbol{t}_c)]$ for the robot interactive control is generated by interpolation. The corresponding impedance $\boldsymbol{\mathcal{K}}_d^m$ and $\boldsymbol{\mathcal{B}}_d^m$ are updated as follows

$$\boldsymbol{\mathcal{K}}_d^m = \boldsymbol{\mathcal{K}}_r + \boldsymbol{\mathcal{K}}_s^{m-1}, \quad \boldsymbol{\mathcal{B}}_d^m = \boldsymbol{\mathcal{B}}_r + \boldsymbol{\mathcal{B}}_s^{m-1}, \tag{7}$$

where $\boldsymbol{\mathcal{K}}_r$ and $\boldsymbol{\mathcal{B}}_r$ are the initial stiffness and damping based on the individual pre-test phase for each human learner and selected reference character. As depicted in Algorithm 1, L initial writing images for the reference character $\boldsymbol{c}$ are collected as $\{\boldsymbol{I}_a^l\}_{l=1}^L$. The initial writing waypoints are extracted from the collected writing images as $\{\{\boldsymbol{\chi}_a^{l,n}\}_{n=1}^N\}_{l=1}^L$. The mean initial writing waypoints $\overline{\boldsymbol{\chi}}_a^I$ can be estimated following Section III-C. $\boldsymbol{\chi}_{\boldsymbol{c}}$ and $\overline{\boldsymbol{\chi}}_a^I$ are aligned with the Dynamic Time Warping (DTW) , as depicted in Fig. 2(d). $\boldsymbol{\mathcal{K}}_r$ is obtained as

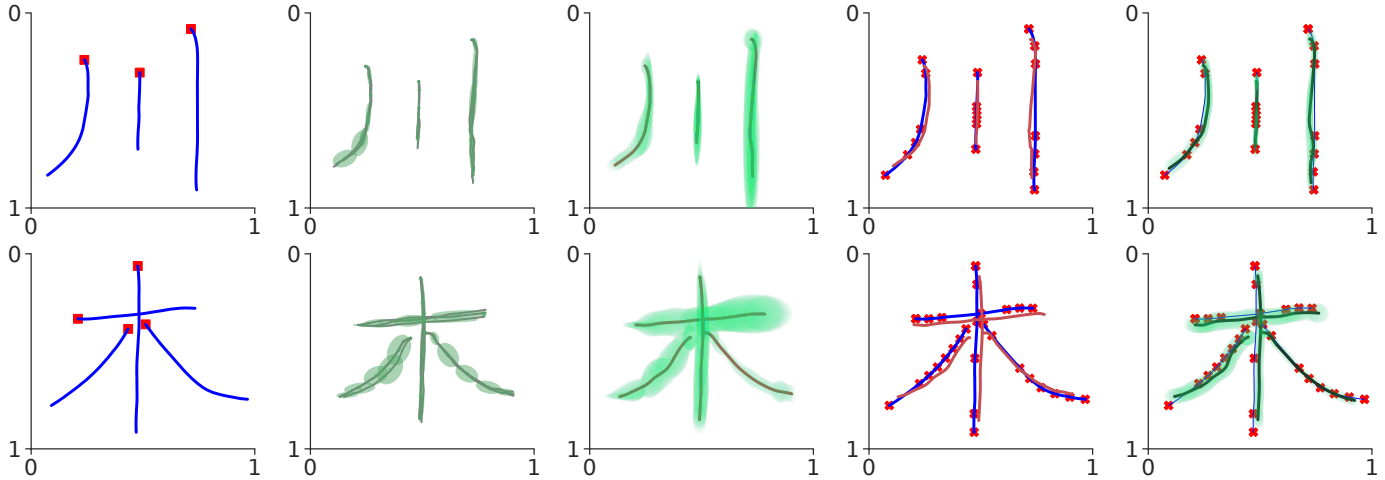
$$\boldsymbol{\mathcal{K}}_r = \beta_r |\Delta \overline{\boldsymbol{\chi}}|, \Delta \overline{\boldsymbol{\chi}} = \overline{\boldsymbol{\chi}}_a - \boldsymbol{\chi}_{\boldsymbol{c}}, \tag{8}$$

where β_r is the coefficient. In practice, damping $\boldsymbol{\mathcal{B}}_r$ is obtained from the stiffness $\boldsymbol{\mathcal{B}}_r = 1/2\sqrt{\boldsymbol{\mathcal{K}}_r}$.

$\boldsymbol{\mathcal{K}}_s^m$ and $\boldsymbol{\mathcal{B}}_s^m$ are the stiffness and damping terms to encourage human engagement. $\boldsymbol{\mathcal{K}}_s^0$ and $\boldsymbol{\mathcal{B}}_s^0$ for the first teaching iteration are set to zero. Then, $\boldsymbol{\mathcal{K}}_s^m$ and $\boldsymbol{\mathcal{B}}_s^m$ are updated according to the $(m-1)$-th writing performance, as follows

$$\begin{aligned} &\boldsymbol{\mathcal{K}}_s^m = \boldsymbol{\mathcal{K}}_s^{m-1} + \beta_{\mathcal{K}} \boldsymbol{\Psi}(\Delta \boldsymbol{\chi}^{m-1}), \boldsymbol{\mathcal{B}}_s^m = 1/2\sqrt{\boldsymbol{\mathcal{K}}_s^m}, \\ &\boldsymbol{\Psi}(\Delta \boldsymbol{\chi}^{m-1,j}) = \frac{\exp\left(\alpha \boldsymbol{\Omega}^{m-1,j} - \Pi_j\right) - 1}{\exp\left(\alpha \boldsymbol{\Omega}^{m-1,j} - \Pi_j\right) + 1}, p = 0,1,2, \end{aligned} \tag{9}$$

where $\boldsymbol{\Psi}(\cdot)$ is an element-wise scalar function (see Fig. 2(g)). $\Delta \boldsymbol{\chi}^{m-1} = \boldsymbol{\chi}^{m-1} - \boldsymbol{\chi}_d^{m-1}$ and $\Delta \boldsymbol{\chi}^{m-1,j}$ is j-th element of $\Delta \boldsymbol{\chi}^{m-1}$. $\boldsymbol{\Omega}^{m-1,j} = \Delta^2 \boldsymbol{\chi}^{m-1,j} - \Pi_j^2$. $\beta_{\mathcal{K}}$ is the coefficient to control the convergence rate, α is a positive scaling scalar. Π_j is the predefined threshold that determines the region whether to increase the assistance to the human learner. In practice, the actual stiffness $\boldsymbol{\mathcal{K}}_d^m$ is bounded by the minimal stiffness $\boldsymbol{\mathcal{K}}_d^{min}$ and the maximal stiffness $\boldsymbol{\mathcal{K}}_d^{max}$, to ensure stable and safe interaction during the robot teaching.



(a) (b) (c) (d) (e)

Fig. 3: Visualization of training waypoints generation for characters C_3^2 and C_4^1. (a) Reference waypoints of each stroke; (b) learned GMMs (green ellipse) from writing waypoints; (c) learned style by GMR from writing waypints; (d) training via-points; (e) generated training waypoints (black lines) by GMR-GP.

TABLE I: Chinese Characters for Teaching

$\mathcal{C}_{s1}$	$\mathcal{C}_{s2}$	$\mathcal{C}_{s3}$	$\mathcal{C}_{s4}$	$\mathcal{C}_{s5}$
七	儿	万	木	丙
刀	又	川	内	兄
入	人	卫	分	冬

TABLE II: Information of Participants

Level #0	Level #1	Level #2
P_0^1 (Indian)	P_1^1 (Singaporean)	P_2^1 (Chinese)
P_0^2 (Indian)	P_1^2 (Singaporean)	P_2^2 (Chinese)
P_0^3 (Turk)	P_1^3 (Malaysian)	P_2^3 (Chinese)
P_0^4 (Ghanaian)	P_1^4 (Malaysian)	P_2^4 (Chinese)
P_0^5 (Korean)	P_1^5 (Indonesian)	P_2^5 (Chinese)

IV. Experiments

We conduct human-subject experiments to address the following questions:

1) Does TeachingBot enable human learners to acquire the target skill more efficiently compared to baseline methods?
2) Can TeachingBot accurately capture individual human writing styles and provide adaptive guidance to human learners?
3) Does the learning experience with TeachingBot encourage active engagement among learners?

In our experiments, participants trained with TeachingBot demonstrated significantly greater improvements in writing quality, particularly in stroke-level similarity to reference characters, compared to baseline methods. The system effectively adapted to individual writing styles, with teaching trajectories progressively aligning with the reference over iterations. Moreover, increased interaction forces reflected greater learner engagement, promoting active participation.

A. Experimental Setups

1) Hardware Setup: We use the 7-DoF torque-controlled Franka Emika Research 3 (FR3) robot for all handwriting teaching experiments. As shown in Fig. 1, the writing workspace measures 350 mm $\times$ 350 mm. Interaction forces at the end-effector handle are estimated from joint torques. The reference character $\boldsymbol{c}$ is represented by a 128 $\times$ 128 image. During pre-test and evaluation, learners' handwritten characters are captured by a mounted camera and resized to 128 $\times$ 128. Reference and writing waypoints are extracted from these images using feature extraction methods. During robotic teaching, the learner's actual writing trajectories are recorded in real time from the end-effector position.

2) Algorithm Setup: The robot teacher's variable impedance control runs at 1000 Hz ($T_s = 0.001$ s). Stiffness ranges from $\boldsymbol{\mathcal{K}}_d^{min} = [200, 200, 200]$ N/m to $\boldsymbol{\mathcal{K}}_d^{max} = [1200, 1200, 1200]$ N/m. Parameters for writing style learning and waypoint generation are $N = 200$, $L = 3$, $H = 5$, and $M = 9$. Impedance variation parameters are set to $\beta_r = 1000$ N/m^2, $\beta_{\mathcal{K}} = 100$ N/m^2, $\Pi_j = 0.05$ m, and $\alpha = 2000$.

3) Training Dataset: The Chinese character dataset $\boldsymbol{\mathcal{C}}$ used in this study is obtained from [34]. We selected 15 distinct characters from the dataset, which includes different numbers of strokes, representing levels of complexity [8]. $\boldsymbol{\mathcal{C}}_{si}$ represents the subdataset including characters with i strokes. $\boldsymbol{c} = C_i^j$ represents the $j-$th character with i strokes. As depicted in Table I, all characters are divided into five groups according to the number of strokes. Each group includes three commonly used Chinese characters, which covers the commonly used strokes in Chinese characters [7, 8, 9]. As depicted in Fig. 3, we visualized the process for generating the training waypoints of characters C_3^2 and C_4^1, following the implementation in Algorithm 1.

4) Participants: 15 subjects (5 females and 10 males) were recruited to participate in the robot teaching experiments. All participants were highly educated university students from

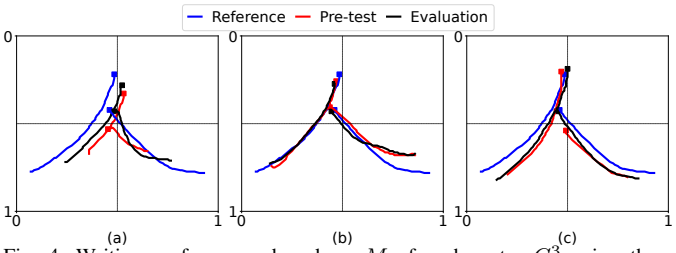


Fig. 4: Writing performance based on M_1 for character C_2^3 using three different teaching methods. The red lines are the writing waypoints from the pre-test phase and the black lines are the writing waypoints from the evaluation phase. (a) FC; (b) RGW; (c) Ours.

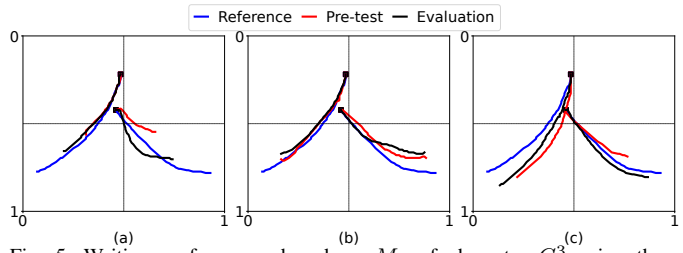


Fig. 5: Writing performance based on M_2 of character C_2^3 using three different teaching methods. The red lines are the writing waypoints from the pre-test phase and the black lines are the writing waypoints from the evaluation phase. (a) FC; (b) RGW; (c) Ours.

diverse national backgrounds. They were categorized into three levels based on their prior knowledge of Chinese character writing, as summarized in Table II. Each participant is denoted as P_i^j, where i represents the level and j represents the subject index within that level.

1) *Level #0*: No prior knowledge or experience with the characters (e.g., Indian, Turkish, Korean students).
2) *Level #1*: Basic familiarity but no regular use (e.g., Singaporean, Malaysian students).
3) *Level #2*: Full proficiency with daily usage (e.g., Chinese students).

B. Experimental Protocol

The effectiveness of TeachingBot is further validated by comparing it to two baselines methods: font copy (FC) and robot guided writing (RGW). For each participant, all 15 Chinese characters in Table I are randomly divided into three groups. Each group of characters include five characters with stroke counts ranging from 1 to 5. Each participant learns three groups of Chinese characters using different methods for the same writing times in a random order, as follows:

1) *TeachingBot*: The group of Chinese characters taught using TeachingBot is referred to as *Group 1* (G_1).
2) *FC*: Participants are only allowed to look at the reference character without any physical correction from the robot. The group of characters taught by FC is referred to as *Group 2* (G_2).
3) *RGW*: Inspired by robotic physical training [17, 14], RGW uses maximum stiffness $\boldsymbol{\mathcal{K}}_d^{max}$ to fully guide participants in replicating the reference trajectory without impedance and trajectory variation. It means that the human learner active engagement is not considered. The group of characters taught by RGW is referred to as *Group 3* (G_3).

The experiments for each character is implemented following three phases:

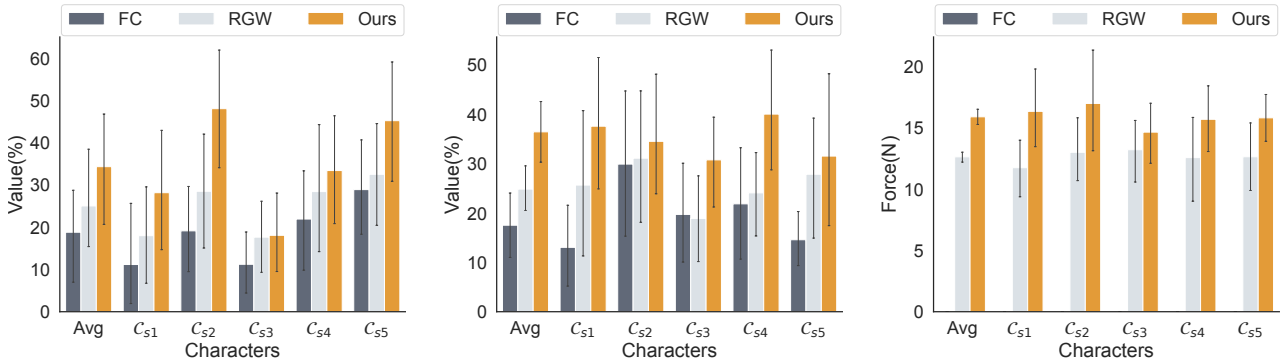
1) *Pre-test*: The pre-test phase is conducted before the teaching phase to assess the initial writing performance of each participant, eliminating the influence of prior knowledge. During this phase, participants write the given reference character L times without any physical or visual guidance. The individual writing duration $\Delta \boldsymbol{t_c}$ for each given reference character is measured for each participant. Then, the initial stiffness and damping for each participant are calculated accordingly.
2) *Teaching*: During the teaching phase, all participants learn characters in G_1 and G_2 by grasping the robot handle and observing the actual writing trajectories. For characters in G_3, all participants first watch a simulated character writing sequence on the screen, after which they attempt to write the character by hand without visual or physical guidance.
3) *Evaluation*: The evaluation phase is conducted after the teaching phase. Participants write the same reference character L times again, without any reference or guidance.

C. Experimental Results

1) Writing Performance: Two key metrics based on Dynamic Time Warping distance are defined to evaluate the similarity between the reference and written waypoints, using two alignment strategies for comparison. First, as shown in Fig. 4, *Metric 1* (M_1) measures global structural similarity by aligning the centers of the written and reference characters, assessing overall shape consistency. Second, as shown in Fig. 5, *Metric 2* (M_2) measures stroke-wise similarity by aligning stroke starting points, focusing on individual stroke accuracy. Finally, participants' active engagement is assessed through the interaction force exerted during the writing task.

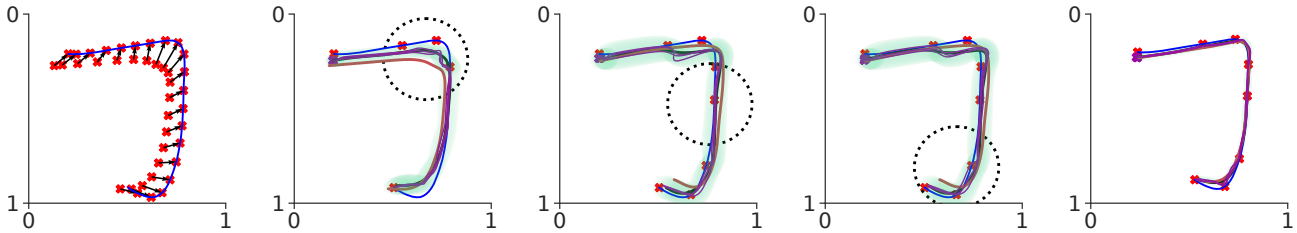
The statistical comparison of writing results for each subdataset $\boldsymbol{\mathcal{C}}_{si}$ is shown in Fig. 6. First, we find that all participants across the three groups have similar initial writing performance. The average percentage improvement in similarity for each group and character after training is shown in Fig. 6(a) and (b). Table III summarizes the average improvement in M_1 and M_2 across characters. Compared to FC and RGW, participants trained with TeachingBot show significantly greater improvement overall. For M_1 (global structural similarity), the improvement compared to the baselines is not entirely significant ($p < 0.05$ vs. FC, $p \approx 0.09$ vs. RGW), suggesting that adapting to individual styles has a marginal effect on global structure. That said, this result still underscores the value of physical guidance in helping learners grasp character structure. In contrast, for M_2 (stroke-wise similarity), TeachingBot shows significantly better improvement than both FC and RGW ($p < 0.05$ for both), highlighting the importance of adaptive strategies in refining fine-grained stroke details.

2) Adaptation of Training Waypoints: To illustrate adaptability during robotic teaching, Fig. 7 shows the training waypoints for character C_2^2 by participant P_2^3. In Fig. 7(a), the red scatter points show the mean writing waypoints from the pre-test, which differ substantially from the reference waypoints (blue line). This difference is used to generate the



(a) (b) (c)

Fig. 6: Statistical results of similarity. Improvement and interaction force of three experimental groups across five groups of Chinese characters. (a) Percentage of similarity improvement based on M_1. (b) Percentage of similarity improvement based on M_2. (c) Average interaction force per waypoint.



(a) (b) (c) (d) (e)

Fig. 7: Visualization of training waypoints changes. Blues lines are reference waypoints, scatter red points are training via-points, red lines are learned means of writing waypoints, black lines are the mean of training waypoints, and purple lines are three sampled potential writing waypoints. (a) Initial reference stiffness generation; (b) 1-st teaching iteration; (c) 3-rd teaching iteration; (d) 6-th teaching iteration; (e) 9-th teaching iteration.

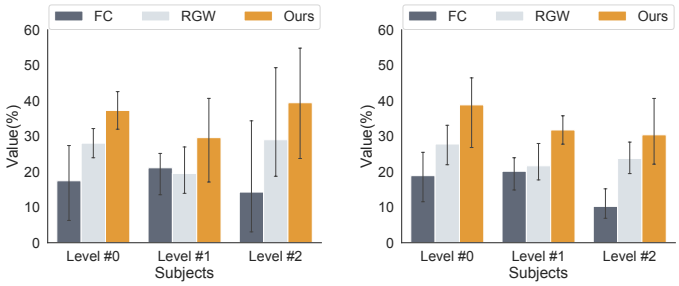
TABLE III: Average Similarity Improvement and Interaction Force.

Method	Value(%)M_1	Value(%)M_2	Force (N)
FC	18.79 $\pm$ 12.15	17.52 $\pm$ 7.62	NA
RGW	25.01 $\pm$ 13.044	24.84 $\pm$ 5.36	12.64 $\pm$ 0.50
Ours	**34.30 $\pm$ 14.03**	**36.44 $\pm$ 6.99**	**15.90 $\pm$ 0.77**

initial reference stiffness. Fig. 7(b)–(e) show the results of the 1st, 3rd, 6th, and 9th teaching iterations, respectively.

In Fig. 7(b), training via-points ($H = 5$) are extracted to enable the human learner to capture the curvature of the reference style in the dashed circle. At 3-rd and 6-th teaching iterations, as illustrated in Fig. 7(c) and (d), training via-points ($H = 8$) are extracted to enable the human learner to focus on different parts of reference style (as shown in the dashed circles). As the robot teaching progresses, the generated training waypoints progressively align with the reference waypoints. Notably, by the 6-th teaching iteration (Fig. 7(d)), the human learner demonstrates significantly reduced variability in replicating the reference character, indicating improved writing precision.

3) Active Engagement of Human Learners: As depicted in Fig. 6(a) and Fig. 6(b), learners using our method show greater similarity improvement compared to those using RGW. To validate the human learner active engagement is crucial for writing skill learning, one indicator of active engagement is the interaction force exerted by the human learner. A higher interaction force suggests the greater active participation during the writing skill learning. We calculate the average interaction force for each character across all learners. As depicted in Fig. 6(c), the interaction force in $G3$ significantly surpasses that in $G2$ ($p < 0.05$), indicating higher human engagement during robot teaching, which aids in improving writing skills. In this paper, we select the same parameters for impedance variation function, which can also be optimized for each learner to achieve higher learning efficiency.



(a) (b)

Fig. 8: Similarity improvement comparison across all characters for participants with varying levels of prior writing experience. (a) Percentage of similarity improvement based on M_1. (b) Percentage of similarity improvement based on M_2.

D. Discussion

We execute post-experiment survey to assess perceived improvement of human learners' in writing skills. Despite $G3$ showing quantitative performance advantages, learners do not report increased confidence in better writing skills compared with $G2$, consistent with previous research [25]. This contradiction leads to two hypotheses: *1*) Lack of visual signal integration: human learners may need clear visual signals in addition to physical ones for better understanding. *2*) Failure to emulate human instructors: due to differences in morphology and the challenge of replicating real teaching scenarios, full emulation is difficult. These hypotheses guide future robot teaching system development, emphasizing the importance of multi-modal signals (e.g., language and visual) for greater effectiveness and generality. In this paper, the effectiveness is mainly validated by highly educated human users. As shown in Fig. 8, human users with different prior writing experiences have tested the performance. The participants with level #2

tend to focus on capturing the global structural similarity, showing improvement in overall similarity when evaluated with metric M_1 compared to M_2. In contrast, participants with level #1 prior experience primarily focus on stroke-wise improvement, exhibiting lower variance in the similarity based on M_2 compared to M_1. These individuals are generally familiar with the structural layout of the characters but lack sufficient practice in writing them fluently. However, no clear trend or consistent conclusion can be drawn for participants with level #0 prior experience, as their performance does not exhibit a distinct pattern. In the future, the human users at different ages will be invited to test the effectiveness. Other characters dataset will be employed to test the effectiveness.

V. Conclusion

In this work, we introduce TeachingBot, a robot teaching system in which robots assume the role of instructors, guiding humans in character writing through adaptive physical interactions. Results from human-subject experiments demonstrate the effectiveness of TeachingBot. Moreover, we provide evidence of how TeachingBot customizes its teaching approach to meet the unique needs of individual learners, leading to enhanced overall engagement and effectiveness. The potential of robot teaching presents exciting opportunities for scaling up education and providing learning opportunities to many, even in the absence of human teachers.

For broader applications of physical robot teaching, our findings suggest that effective teaching of physical skills requires personalized guidance, dynamic adjustment of assistance levels, and the integration of physical guidance with multimodal feedback such as visual or verbal cues.

References

[1] D. P. Losey, C. G. McDonald, E. Battaglia, and M. K. O'Malley, "A review of intent detection, arbitration, and communication aspects of shared control for physical human–robot interaction," *Applied Mechanics Reviews*, 2018.

[2] M. Arduengo, A. Colomé, J. Borràs, L. Sentis, and C. Torras, "Task-adaptive robot learning from demonstration with gaussian process models under replication," *IEEE Robotics and Automation Letters*, 2021.

[3] Y. Zhou, J. Gao, and T. Asfour, "Learning via-point movement primitives with inter-and extrapolation capabilities," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.

[4] D. P. Losey, H. J. Jeon, M. Li, K. Srinivasan, A. Mandlekar, A. Garg, J. Bohg, and D. Sadigh, "Learning latent actions to control assistive robots," *Autonomous robots*, vol. 46, no. 1, pp. 115–147, 2022.

[5] A. U. Pehlivan, D. P. Losey, and M. K. O'Malley, "Minimal assist-as-needed controller for upper limb robotic rehabilitation," *IEEE Transactions on Robotics*, 2015.

[6] C. Yu, Y. Xu, L. Li, and D. Hsu, "Coach: Cooperative robot teaching," in *Proceedings of The 6th Conference on Robot Learning*, 2023.

[7] S. Lemaignan, A. Jacq, D. Hood, F. Garcia, A. Paiva, and P. Dillenbourg, "Learning by teaching a robot: The case of handwriting," *IEEE Robotics & Automation Magazine*, vol. 23, no. 2, pp. 56–66, 2016.

[8] Y. Luo, Z. Wu, and Z. Lian, "Callirewrite: Recovering handwriting behaviors from calligraphy images without supervision," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 8671–8678.

[9] X. Wang and Z. Gong, "Rodal: style generation in robot calligraphy with deep adversarial learning," *Applied Intelligence*, vol. 54, no. 17, pp. 7913–7923, 2024.

[10] Y. Huang, J. Silvério, L. Rozo, and D. G. Caldwell, "Generalized task-parameterized skill learning," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 5667–5474.

[11] Y. Huang, L. Rozo, J. Silvério, and D. G. Caldwell, "Kernelized movement primitives," *The International Journal of Robotics Research*, 2019.

[12] N. Jaquier, D. Ginsbourger, and S. Calinon, "Learning from demonstration with model-based gaussian process," in *Conference on Robot Learning*, 2020.

[13] T. Proietti, V. Crocher, A. Roby-Brami, and N. Jarrasse, "Upper-limb robotic exoskeletons for neurorehabilitation: a review on control strategies," *IEEE reviews in biomedical engineering*, vol. 9, pp. 4–14, 2016.

[14] Z. Yang, S. Guo, Y. Liu, M. Kawanishi, and H. Hirata, "A task performance-based semg-driven variable stiffness control strategy for upper limb bilateral rehabilitation system," *IEEE/ASME Transactions on Mechatronics*, 2022.

[15] Z. Hou, T. Ma, W. Wang, and H. Yu, "Contextual policy search for task-level adaptation in physical human–robot interaction," *IEEE/ASME Transactions on Mechatronics*, 2025.

[16] S. Pareek, H. J. Nisar, and T. Kesavadas, "Ar3n: A reinforcement learning-based assist-as-needed controller for robotic rehabilitation," *IEEE Robotics & Automation Magazine*, 2023.

[17] X. Li, Y.-H. Liu, and H. Yu, "Iterative learning impedance control for rehabilitation robots driven by series elastic actuators," *Automatica*, 2018.

[18] M. Srivastava, E. Biyik, S. Mirchandani, N. Goodman, and D. Sadigh, "Assistive teaching of motor control tasks to humans," in *Advances in Neural Information Processing Systems*, 2022.

[19] A. Singla, I. Bogunovic, G. Bartók, A. Karbasi, and A. Krause, "Near-optimally teaching the crowd to classify," *International Conference on Machine Learning*, 2014.

[20] S. Zilles, S. Lange, R. Holte, and M. Zinkevich, "Models of cooperative teaching and learning," *Journal of Machine Learning Research*, 2011.

[21] T. Doliwa, H. U. Simon, and S. Zilles, "Recursive teaching dimension, learning complexity, and maximum classes," pp. 209–223, 2010.

[22] O. Mac Aodha, S. Su, Y. Chen, P. Perona, and Y. Yue, "Teaching categories to human learners with visual explanations," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[23] M. Srivastava, N. Goodman, and D. Sadigh, "Generating language corrections for teaching physical control tasks," in *40th International Conference on Machine Learning (ICML)*, 2023.

[24] D. Hood, S. Lemaignan, and P. Dillenbourg, "When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting," in *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, 2015, pp. 83–90.

[25] R. Tian, M. Tomizuka, A. D. Dragan, and A. V. Bajcsy, "Towards modeling and influencing the dynamics of human learning," *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023.

[26] A. N. Rafferty, E. Brunskill, T. L. Griffiths, and P. Shafto, "Faster teaching by pomdp planning," in *Artificial Intelligence in Education*, 2011.

[27] A. Sandygulova, W. Johal, Z. Zhexenova, B. Tleubayev, A. Zhanatkyzy, A. Turarova, T. Asselborn, and P. Dillenbourg, "Cowriting kazakh: Learning a new script with a robot," in *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2020, pp. 113–120.

[28] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, "Learning and generalization of motor skills by learning from demonstration," in *2009 IEEE International Conference on Robotics and*

Automation, 2009.

[29] A. Paraschos, C. Daniel, J. Peters, and G. Neumann, "Using probabilistic movement primitives in robotics," *Autonomous Robots*, 2018.

[30] M. Maaref, A. Rezazadeh, K. Shamaei, R. Ocampo, and T. Mahdi, "A bicycle cranking model for assist-as-needed robotic rehabilitation therapy using learning from demonstration," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 653–660, 2016.

[31] Z. Warraich and J. A. Kleim, "Neural plasticity: the biological substrate for neurorehabilitation," *Pm&r*, 2010.

[32] E. Caldarelli, A. Colomé, and C. Torras, "Perturbation-based stiffness inference in variable impedance control," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8823–8830, 2022.

[33] Z. Hou, W. Yang, R. Chen, P. Feng, and J. Xu, "A hierarchical compliance-based contextual policy search for robotic manipulation tasks with multiple objectives," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 4, pp. 5444–5455, 2022.

[34] S. Kishore, "Make me a hanzi demo," https://github.com/skishore/makemeahanzi.