IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

TeachingBot: Robot Teacher for Human Handwriting

Zhimin Hou*, Cunjun Yu*, David Hsu *Fellow IEEE*, Haoyong Yu *Senior Member, IEEE*

***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 human learner and (ii) maintaining an engaging learning experience. For the first challenge, TeachingBot uses a probabilistic model to capture the human 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 human learner's natural writing. For the second challenge, TeachingBot employs variable impedance control to guide the human learner, dynamically adjusting the strength of physical guidance based on the human learner's writing performance. Human-subject experiments demonstrate the effectiveness of TeachingBot, showing clear improvement in learners' handwriting and engagement over baseline methods.**

***Index Terms*—Robot-assisted teaching, Physical human-robot interaction**

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]. 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) leverage individual writing styles to generate effective gradual guidance pathways toward standard character forms and (2) provide an appropriate level of physical guidance that promotes active engagement.

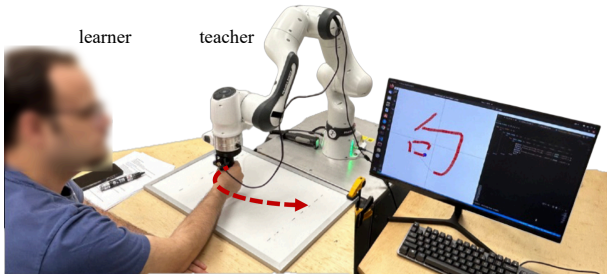

Fig. 1: *TeachingBot*: the robot teacher guides the human learner in writing Chinese characters, adapting to the human learner's preferences and proficiency.

We present ***TeachingBot***, an adaptive robot teaching system designed to teach Chinese character writing through physical guidance. The human learner holds the end effector of the robotic arm, which guides them through the generated teaching trajectory and compliant physical interaction control (see the illustration in Fig. 1). Our key insight is that effective teaching should use the human learner's current writing style as a starting point to generate gradual reference trajectories rather than immediately imposing reference strokes. TeachingBot creates a personalized learning pathway that feels natural to the human learner while systematically guiding them towards the reference character. TeachingBot achieves this through a two-step adaptive approach to encourage human learners to progressively converge toward standard reference character forms.

First, the system captures the human learner's writing style by collecting their writing trajectories and modeling them probabilistically. We combine Gaussian Mixture Regression (GMR) with Gaussian Process (GP) regression, where the GMR model learned from the human learner's writing serves as the prior for the GP. This allows us to encode individual variability directly into trajectory generation. Unlike existing learning from demonstration based methods [10, 11, 12] that encode expert demonstrations for fixed reference trajectories, our method generates intermediate teaching trajectories that progressively adapt from each human learner's current style

Manuscript received: July 7, 2025; Revised: October 9, 2025; Accepted: December 7, 2025. This paper was recommended for publication by Editor Sonia Chernova upon evaluation of the Associate Editor and Reviewers' comments. This research is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG2-PhD-2022-01-036[T]) and in part by the Ministry of Education, Singapore under the Academic Research Fund grant (T1 251RES2406).

*Zhimin Hou and Cunjun Yu contributed equally to this work.

Zhimin Hou is with School of Data Science, Lingnan University, Hong Kong, and with the Department of Biomedical Engineering, National University of Singapore, Singapore 117583.

Cunjun Yu is with School of Computing, National University of Singapore.

David Hsu is with School of Computing, Smart System Institute, National University of Singapore.

Haoyong Yu is with the Department of Biomedical Engineering, National University of Singapore, Singapore 117583.

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

Corresponding author: Cunjun Yu (e-mail: cunjun.yu@u.nus.edu).

Digital Object Identifier (DOI): see top of this page.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

toward the standard target, encouraging human learners' active engagement and mainly capturing the key via-points.

Second, the robot executes the teaching trajectory through compliant physical interaction control [13, 14, 15]. We develop a variable impedance controller (VIC) to dynamically adjust the guidance force by optimizing the impedance parameters: applying stronger corrective assistance when the human learner deviates significantly from the reference, and reducing assistance when deviations are minor to encourage exploration and sustain engagement [5, 16, 17]. Unlike existing VIC-based approaches for rehabilitation, our method first derives a greedy learner-adaptive initial stiffness to maximize active engagement and then iteratively updates the stiffness to reinforce corrective guidance for each personalized reference trajectory.

Through human-subject experiments with 30 participants with varying levels of Chinese writing proficiency, we demonstrate that TeachingBot significantly improves handwriting quality compared to three baseline methods. Participants showed greater accuracy in replicating both the overall structure and fine-grained stroke details of characters. Furthermore, increased interaction forces during training indicate improved learner engagement, highlighting the effectiveness of adaptive physical guidance in promoting active skill acquisition.

II. Related Work

A. Robot-Assisted Motor Skill Acquisition in pHRI

Physical human–robot interaction (pHRI) has proven effective for motor skill acquisition [18, 7, 6], with direct physical guidance supporting fine motor control more effectively than social interaction alone [15, 19]. Most robot-assisted training research has focused on motor rehabilitation [13, 5, 20], where the primary objective is restoring normative motor functions in individuals with physical or neurological impairments. Robotic assistive strategies provide adaptive guidance to facilitate functional movement recovery. Assist-as-needed (AAN) controllers have demonstrated strong potential for promoting active patient engagement [5, 20, 21], with their efficacy validated through laboratory and clinical studies.

Our work addresses an underexplored domain: handwriting skill acquisition. Unlike rehabilitation, which primarily targets motor control, handwriting learning involves both motor and cognitive processes. Human learners possess intact motor functions but must develop cognitive understanding of character structure, stroke order, and spatiotemporal organization. In rehabilitation, robotic assistance compensates for motor impairments to achieve fixed functional targets. In contrast, handwriting training requires learners with diverse stylistic baselines to internalize culturally defined writing standards through progressive cognitive–motor adaptation. This distinction necessitates different strategies: robots cannot simply guide learners along fixed expert trajectories; they must generate progressive intermediate targets that facilitate cognitive processing and gradual internalization of correct patterns.

B. Reference Trajectory Generation in pHRI

Reference trajectory generation is a core component of pHRI control. Probabilistic methods, including Gaussian Mixture Models/Regression (GMM/GMR) and Dynamic Movement Primitives (DMPs), are commonly used to reproduce smooth and human-like trajectories from expert demonstrations [10, 11, 22]. Extensions such as Probabilistic DMP (ProMP) and GP models enhance adaptability by encoding trajectory families through basis functions [23, 12, 3]. Optimization-based methods partially deform reference trajectories according to specific task targets or individual biomechanical conditions [11, 15]. Most existing probabilistic methods encode *therapist-informed demonstrations* for safe physical training [24], requiring re-demonstration when task targets change. Partial trajectory deformation for patient-specific requirements has been validated in upper-limb bilateral training [15] and lower-limb rehabilitation [21].

However, few studies have updated reference trajectory for progressive skill learning. This work extends AAN training to the trajectory level for handwriting skill acquisition. Rather than relying on expert demonstrations, our method encodes the human learner's current variability to generate intermediate reference trajectories that progressively bridge the gap between the human learner's current writing style and the desired character form. This approach allows human learners to preserve individual motion preferences and actively engage while gradually converging toward key features.

C. Variable Impedance Control in pHRI

Physical guidance effectiveness depends critically on how assistance intensity is modulated [13, 25]. Active user engagement promotes neural plasticity [26, 5] and enhances motor skill learning efficiency [19]. Among various strategies, VIC-based methods have shown strong potential for AAN training. Yang et al. [14] developed tracking-error-based VIC for upper-limb rehabilitation, where assistance adapts to patient performance. Higher impedance increases physical guidance, while lower impedance encourages active participation through minimal assistance [17]. Pehlivan et al. [5, 20] regulated assistance levels around predefined trajectories using estimated human output force from external sensors or observers. Learning-based methods adapt impedance parameters from accumulated interaction experience for fixed task targets [16, 17, 27].

However, existing methods often require parameters manually tuning or individualized estimation for each learner. Handwriting is a highly dynamic motor learning process with fine-grained temporal and spatial variations. Our VIC scheme introduces two key adaptations for robot-aided teaching: *1)* learner-specific stiffness derived directly from initial writing performance to maximize active engagement; and *2)* iteratively update stiffness to progressively reinforce key via-point capture while maintaining engagement as the remaining trajectory aligns with the personalized reference.

III. Methodology

A. Problem Formulation

We build on prior work [6, 28] 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. The objective of TeachingBot is to adaptively teach human learners to write

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

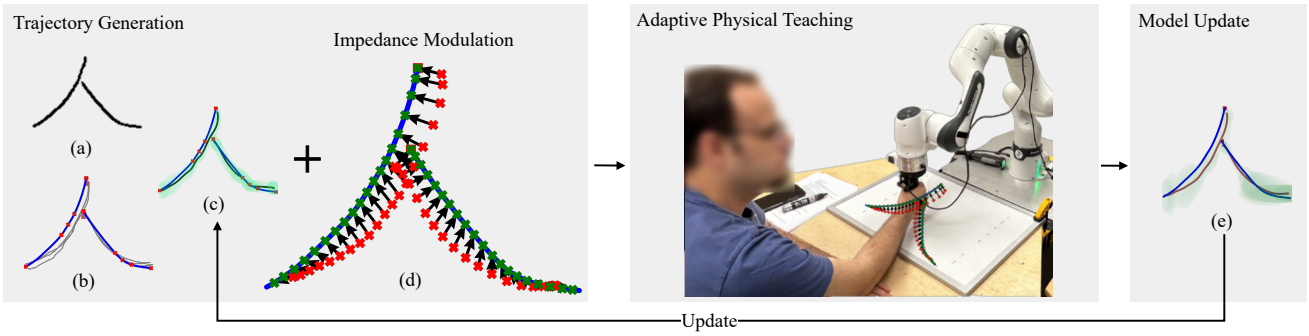

Fig. 2: Overview of TeachingBot. First, a reference character (a) is selected, and waypoints (b) are extracted from both the reference and the learner's handwriting. Their differences guide initial impedance modulation (d), where the blue/green path shows the reference trajectory/waypoints, the red cross 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).

the reference character. 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 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 VIC $\pi(\boldsymbol{u}|\boldsymbol{x},\dot{\boldsymbol{x}};\boldsymbol{x}_d,\dot{\boldsymbol{x}}_d,\boldsymbol{\mathcal{K}}_d,\boldsymbol{\mathcal{B}}_d)$ is designed 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. $\boldsymbol{\mathcal{K}}_d \in \mathbb{R}^{3\times3}$ and $\boldsymbol{\mathcal{B}}_d \in \mathbb{R}^{3\times3}$ are the reference stiffness and damping. 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). $\boldsymbol{\mathcal{K}}_d$ and $\boldsymbol{\mathcal{B}}_d$ should be regulated to adjust the assistance level provided by the *robot teacher* [13, 14, 17].

B. Overview of TeachingBot

The overview of TeachingBot is depicted in Fig. 2. The reference character is denoted by the variable $\boldsymbol{c} \in \mathbf{C}$. $\boldsymbol{s}_i \in \mathbf{S_c}$ represents i-th stroke of character $\boldsymbol{c}$. The dataset of reference Chinese characters, $\mathbf{C}$, includes images of each character $\{\boldsymbol{I}_c\}$ for the human learners to learn (see Fig. 2(a)). $\mathbf{S_c}$ 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 an image processing method, N is the number of waypoints and $\boldsymbol{\chi}_c^n$ denotes the n-th reference waypoint. As illustrated in Fig. 2(b), the reference waypoints of two strokes are extracted and plotted as the blue lines.

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, $\boldsymbol{\mathcal{D}}_L^m = \{\{\boldsymbol{t}_n^l,\ \boldsymbol{\chi}_a^{l,n}\}_{n=1}^N\}_{l=1}^L$, 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 $\boldsymbol{\mathcal{D}}_V^m$, defined as training via-points. A GMR-GP model, $f(\boldsymbol{\chi}_r|\boldsymbol{t},\boldsymbol{\mathcal{D}}_L^m,\boldsymbol{\mathcal{D}}_V^m)$, 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 end-effector, 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 $\boldsymbol{\mathcal{D}}_L^{m-1}$[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).

Algorithm 1 TeachingBot

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

D. Training Via-points Extraction

The curvature-based trajectory compression is employed to extract the training via-points from the reference waypoints $\boldsymbol{\chi}_c$. We retain waypoints where the curvature (change in direction) is highest, as these waypoints are likely to represent features of the reference character [29]. 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. During the m-th teaching iteration, the training via-points are obtained and stored in $\boldsymbol{\mathcal{D}}_V^m = \{\boldsymbol{t}_h^m, \boldsymbol{\chi}_c^h\}_{h=1}^H$ ($H << N$). The corresponding time for each training via-point is estimated by aligning with the collected actual writing waypoints using same writing duration. 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 $\boldsymbol{\mathcal{D}}_L^{m-1}$, the training waypoints need to be generated to maintain the current writing style and follow the instruction of training waypoints in $\boldsymbol{\mathcal{D}}_V^m$. 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 $\boldsymbol{\xi}_o = \boldsymbol{\chi}$, $\boldsymbol{\epsilon}_t = [\epsilon_t^1, \epsilon_t^p, \cdots, \epsilon_t^{d_o}]$, $\epsilon_t^p \sim \mathcal{N}(0, \sigma_p^2)$. The distribution of $\boldsymbol{\chi}$ input $\boldsymbol{t}$ is given by $\boldsymbol{\chi}(\boldsymbol{t}) \sim \boldsymbol{\mathcal{GP}}(\boldsymbol{\mu}(\boldsymbol{t}), \boldsymbol{k}(\boldsymbol{t}, \boldsymbol{t}'))$, 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 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 $\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \boldsymbol{\mathcal{D}}_L^{m-1}) \sim \mathcal{N}(\boldsymbol{\mu}_L^*, \boldsymbol{\Sigma}_L^*)$. 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 $\boldsymbol{\mathcal{D}}_V^m$ are treated as new observations to complement the previous observations in $\boldsymbol{\mathcal{D}}_L^{m-1}$. Unlike standard MOGP that uses a fixed prior mean (typically zero or constant), our GMR-GP model replaces $\boldsymbol{\mu}(\boldsymbol{t}_*)$ with the GMR mean estimated from the human learner's writing style, encoding their individual variability as the prior. This allows the posterior to generate intermediate training waypoints that naturally bridge the gap between the human learner's current style and the expert target, enabling smooth transitions that promote engagement rather than abrupt corrections that may discourage learners. 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

$$\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \boldsymbol{\mathcal{D}}_L^{m-1}, \boldsymbol{\mathcal{D}}_V^m) \propto \mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \boldsymbol{\mathcal{D}}_L^{m-1}) \mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \boldsymbol{\mathcal{D}}_V^m), \tag{6}$$

where $\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \boldsymbol{\mathcal{D}}_L^{m-1}, \boldsymbol{\mathcal{D}}_V^m) \sim \mathcal{N}(\boldsymbol{\mu}_{L-V}^m, \boldsymbol{\Sigma}_{L-V}^m)$, $\boldsymbol{\mu}_{L-V}^m$ 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 $\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \boldsymbol{\mathcal{D}}_L^{m-1})$; $\boldsymbol{\mu}_V^*$ and $\boldsymbol{\Sigma}_V^*$ are the mean and covariance of $\mathcal{P}(\boldsymbol{\chi}^*|\boldsymbol{t}_*, \boldsymbol{\mathcal{D}}_V^m)$. 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 $\boldsymbol{c}$, which will be estimated from each human learner during the pre-test phase. After obtaining the training waypoints $\{\boldsymbol{\chi}_d^{m,n}\}_{n=1}^N$ from the m-th teaching iteration, the reference trajectory $\boldsymbol{\mathcal{T}}_d =$

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

$[\boldsymbol{x}_d(0), \boldsymbol{x}_d(T_s), \cdots, \boldsymbol{x}_d(\Delta \boldsymbol{t}_c)]$ 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. In summary, 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). The initial stiffness is set proportional to the human learner's initial deviation from the reference:

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

where β_r is the coefficient. This *greedy* initialization provides individualized assistance from the first iteration without requiring external force sensors, maximizing engagement by matching the initial assistance level to each learner's starting capability. 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

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

where $\boldsymbol{\Psi}(\cdot)$ is an element-wise scalar function. $\Delta \boldsymbol{\chi}^{m-1} = \boldsymbol{\chi}^{m-1} - \boldsymbol{\chi}_d^{m-1}$ 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.

IV. Experiments

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

1) Does TeachingBot accelerate skill acquisition compared to baselines?
2) Does TeachingBot capture individual writing styles and provide adaptive guidance?
3) Does TeachingBot promote active learner engagement?

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 × 350 mm.

TABLE I: Chinese Characters for Teaching

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

2) Algorithm Setup: Our VIC for robot teacher 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: We selected 15 Chinese characters from [30], spanning 1–5 strokes to represent varying complexity [8]. $\mathbf{C}_{si}$ denotes characters with i strokes; $\boldsymbol{c} = C_i^j$ denotes the j-th character with i strokes. Characters are grouped by stroke count (Table I). Fig. 3 visualizes training waypoint generation for C_3^2 and C_4^1.

4) Participants: 30 subjects were recruited to participate in the robot teaching experiments. All participants were highly educated university students from diverse national backgrounds. They were categorized into three levels based on their prior knowledge of Chinese character writing, as follows:

- *Level #0*: No prior knowledge (e.g., Indian, Turk).
- *Level #1*: Basic familiarity, no regular use (e.g., Singaporean, Malaysian).
- *Level #2*: Full proficiency, daily usage (e.g., Chinese).

B. Experimental Protocol

The effectiveness of TeachingBot is further validated by comparing it to three baselines methods: font copy (FC), robot guided writing (RGW), and variable impedance control with weighted trajectory combination (VIC-WTC). For each participant in the first 15 participants, 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 users and characters taught using TeachingBot is referred to as *Group 1* (G_1).
2) *FC*: Participants are only allowed to look at the reference character and then practice on their own without any physical correction from the robot. The group of users and 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 users and characters taught by RGW is referred to as *Group 3* (G_3).

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

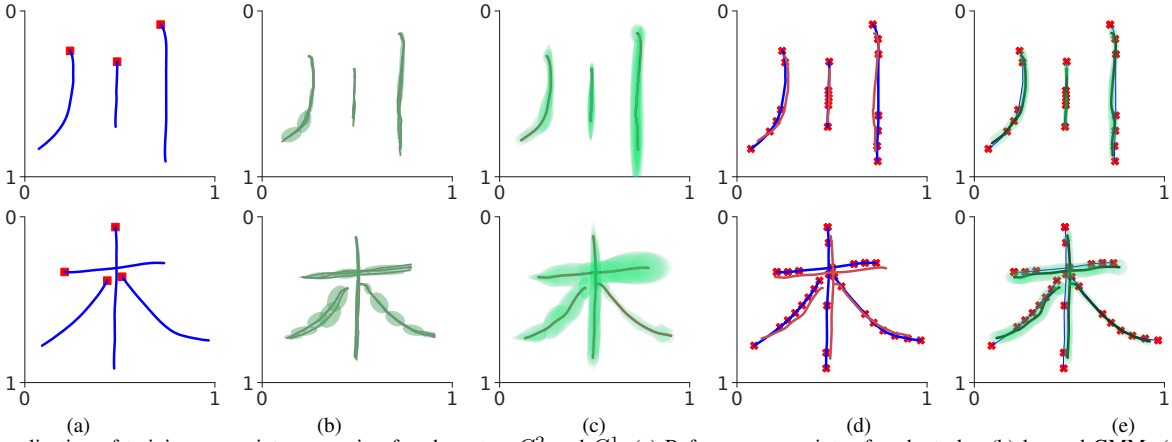

Fig. 3: Visualization of training waypoints generation for characters C_3^2 and C_4^1. (a) Reference waypoints of each stroke; (b) learned GMMs (green ellipse) from writing waypoints; (c) learned style by GMR from writing waypints; (d) training via-points; (e) generated training waypoints (black lines)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

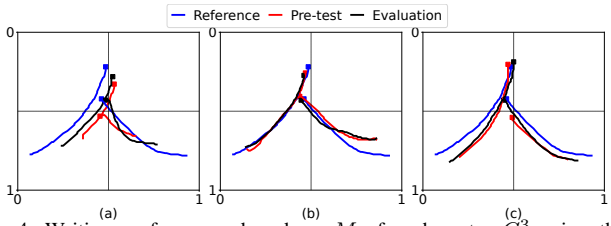

Fig. 4: Writing performance based on M_1 for character C_2^3 using three different teaching methods. The red lines are the writing waypoints from the pre-test phase and the black lines are the writing waypoints from the evaluation phase. (a) FC; (b) RGW; (c) Ours.

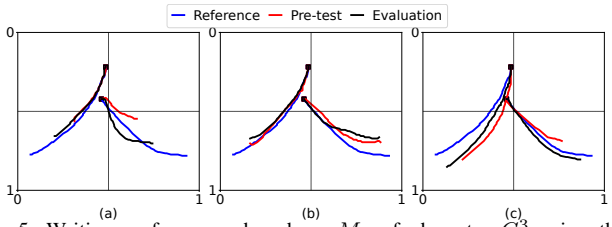

Fig. 5: Writing performance based on M_2 of character C_2^3 using three different teaching methods. The red lines are the writing waypoints from the pre-test phase and the black lines are the writing waypoints from the evaluation phase. (a) FC; (b) RGW; (c) Ours.

A separate between-subject experiment with 15 additional participants of varying skill levels evaluates generalizability using the VIC-WTC baseline. VIC-WTC provides variable assistance through weighted blending of user and reference trajectories. This group is denoted as *Group 4* (G_4). Each character experiment follows three phases:

1) *Pre-test:* Participants write each reference character L times without guidance to establish baseline performance. Initial writing speed $\Delta \boldsymbol{t_c}$ is measured; stiffness and damping parameters are computed.
2) *Teaching:* For G_1 and G_2, participants grasp the robot handle and observe trajectories. For G_3, participants watch a simulated writing sequence, then reproduce the character independently.
3) *Evaluation:* Participants write each reference character L times without guidance to assess performance.

C. Experimental Results

1) Writing Performance: We define two metrics based on DTW distance to evaluate similarity between reference and written waypoints. *Metric 1* (M_1) measures global structural similarity by aligning character centers, assessing overall shape consistency (Fig. 4). *Metric 2* (M_2) measures stroke-wise similarity by aligning stroke starting points, focusing on individual stroke accuracy (Fig. 5). Active engagement is assessed through interaction force exerted during writing. The statistical comparison of writing results for each subdataset $\mathbf{C}_{si}$ 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). Compared to baselines, participants trained with TeachingBot show greater improvement overall. For M_1 (global structural similarity), the improvement compared to baselines is not entirely significant ($p < 0.05$ vs. FC, $p \approx 0.09$ vs. RGW, $p > 0.1$ vs. VIC-WTC), suggesting that variable impedance control and physical guidance primarily contributes to helping learners grasp overall character structure. However, for M_2 (stroke-wise similarity), TeachingBot shows significantly better improvement than all baselines ($p < 0.05$ vs. FC, RGW, and VIC-WTC), demonstrating that the GMR-GP based adaptive trajectory deformation is essential for capturing fine-grained stroke details and refining individual stroke accuracy.

2) Adaptation of Training Waypoints: Fig. 7 illustrates adaptability for participant P_2^3. In (a), red scatter points (pre-test mean waypoints) differ substantially from reference waypoints (blue line). This difference initializes reference stiffness. Panels (b)–(e) show iterations 1, 3, 6, and 9. In (b), training via-points ($H = 5$) guide the human learner to capture curvature (dashed circle). By iterations 3 and 6 (c, d), via-points ($H = 8$) target different reference features (dashed circles). Training waypoints progressively align with the reference. By iteration 6 (d), variability in replicating the reference character is significantly reduced, confirming improved 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. One indicator of active engagement is the interaction force exerted by the human learner. A higher interaction force suggests 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 both G_1 and G_4 significantly surpasses that in G_3 ($p < 0.05$), demonstrating that VIC is essential for promoting active human engagement during robot teaching. In this paper, we select the same parameters for impedance variation function, which can also be optimized

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

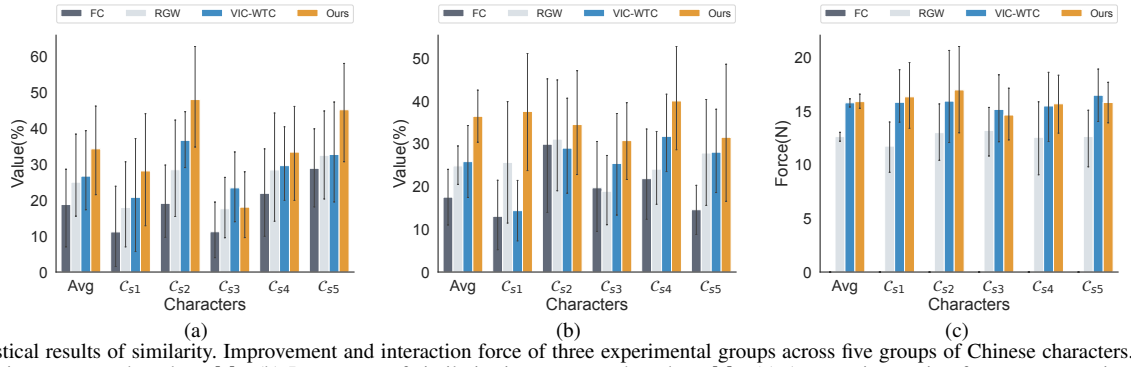

Fig. 6: Statistical results of similarity. Improvement and interaction force of three experimental groups across five groups of Chinese characters. (a) Percentage of similarity improvement based on M_1. (b) Percentage of similarity improvement based on M_2. (c) Average interaction force per waypoint.

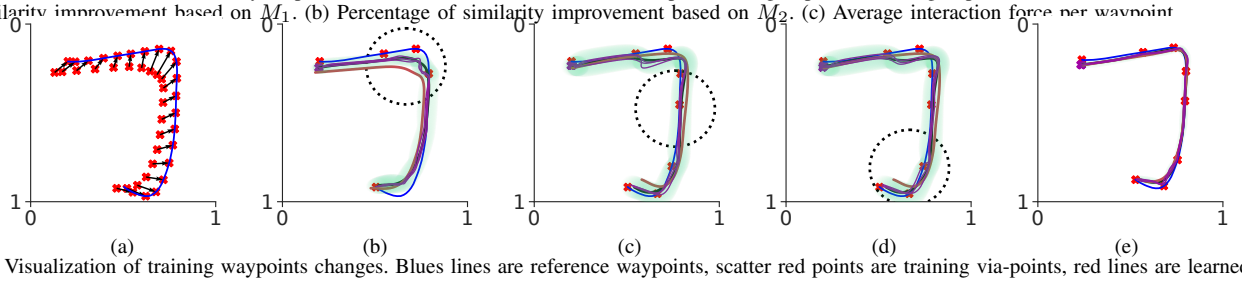

Fig. 7: Visualization of training waypoints changes. Blues lines are reference waypoints, scatter red points are training via-points, red lines are learned means of writing waypoints, black lines are the mean of training waypoints, and purple lines are three sampled potential writing waypoints. (a) Initial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

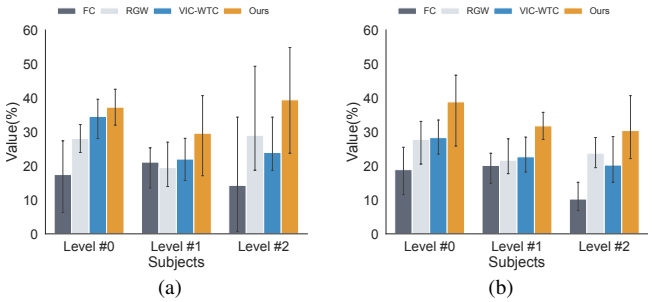

Fig. 8: Similarity improvement comparison of all characters for participants with varying levels of prior writing experience. (a) Percentage of similarity improvement using M_1. (b) Percentage of similarity improvement using M_2.

for each learner to achieve higher learning efficiency.

D. Discussion

We conducted a post-experiment survey to assess perceived improvement of human learners' in writing skills. Despite G_3 showing quantitative performance advantages, learners do not report increased confidence in better writing skills compared with G_2, consistent with previous research [31]. 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 emphasize the importance of multi-modal signals (e.g., language and visual) for greater effectiveness and generality. This work validated its effectiveness by highly educated human users. The results in Fig. 8 indicate that participants with *Level #2* tend to capturing the global structural similarity, showing improvement in overall similarity when evaluated with metric M_1 compared to M_2. 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.

Our method was experimentally validated against three baseline approaches. FC represents an ideal zero-impedance control scheme in which the robot provides no physical assistance. RGW corresponds to an VIC-based method with maximum stiffness. VIC-WTC adopts the same variable impedance control scheme as our method but updates the reference trajectory using a weight combination scheme. For fair comparison, all methods display the visual reference for the human learners so that the writing error will not be reinforced during training. The results in Fig. 6 and Fig. 8 indicate that both reference trajectory deformation and VIC-based adaptive interaction are essential for achieving efficient robot teaching. While VIC-WTC assigns equal weights to the deformation of all waypoints, our GMR-GP model generates a reference trajectory that adaptively emphasizes key via-points. By incorporating this adaptive reference trajectory deformation, our method effectively mitigates the large initial interaction forces commonly observed at the beginning of the teaching process. Moreover, a learner-specific adaptive stiffness is greedily predefined based on the human learner's initial writing performance to maximize active participation. The impedance parameters are further iteratively updated during each writing iteration to reinforce the human learner's engagement to capture and reproduce key via-points.

Three limitations remain. First, this work demonstrates feasibility. Future work will evaluate effectiveness with more learners of diverse ages and an expanded character set. Second, new evaluation metrics (e.g., progressive scoring [19]) maybe better integrate the human learner's writing style with reference targets. Third, the time-dependent GMR-GP model indicates stroke order but relies on estimated average writing duration per teaching iteration. Future work will explore alternatives to accommodate speed variation.

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 learners, 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.

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] 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.

[19] D. F. P. Granados, B. A. Yamamoto, H. Kamide, J. Kinugawa, and K. Kosuge, "Dance teaching by a robot: Combining cognitive and physical human–robot interaction for supporting the skill learning process," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1452–1459, 2017.

[20] A. U. Pehlivan, D. P. Losey, C. G. Rose, and M. K. O'Malley, "Maintaining subject engagement during robotic rehabilitation with a minimal assist-as-needed (maan) controller," in *2017 International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 2017, pp. 62–67.

[21] Z. Yang, H. Jin, W. Gao, E. Wang, Y. Shu, M. Wu, and S. Zhang, "Hierarchical trajectory deformation algorithm with hybrid controller for active lower limb rehabilitation," *IEEE Robotics and Automation Letters*, vol. 9, no. 7, pp. 6240–6247, 2024.

[22] 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.

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

[24] Y. Zhang, T. Li, H. Tao, F. Liu, B. Hu, M. Wu, and H. Yu, "Research on adaptive impedance control technology of upper limb rehabilitation robot based on impedance parameter pblackiction," *Frontiers in Bioengineering and Biotechnology*, vol. 11, p. 1332689, 2024.

[25] J. L. Emken, R. Benitez, and D. J. Reinkensmeyer, "Human-robot cooperative movement training: learning a novel sensory motor transformation during walking with robotic assistance-as-needed," *Journal of neuroengineering and rehabilitation*, vol. 4, no. 1, p. 8, 2007.

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

[27] 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.

[28] 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.

[29] H. Asada and M. Brady, "The curvature primal sketch," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 1, pp. 2–14, 1986.

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

[31] 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.