2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

A Novel Human-Machine Dual-Task Gaming Framework for Visual-Attention Training

Fengjun Mu[1], Jingting Zhang[1*], Zonghai Huang[1], Chen Chen[1], Chaobin Zou[1], Guangkui Song[1], Hong Cheng[1]

Abstract— **Efficient brain functional training with rehabilitation robots has been an important and challenging topic in the human-machine interaction (HMI) field. Adjusting the interaction and gaming behaviors between human and machine to effectively activate the brain's functional behavior is still a substantial challenge. In this paper, we take the visual-attention training as an example, and propose a novel human-machine co-gaming interaction framework by integrating a dual-task gaming paradigm and a human–machine gaming strategy. It has a remarkable capability of effectively utilizing the gaming characteristics of HMI behaviors and tasks, to effectively and precisely activate the human's active attention and passive attention for training. Specifically, we design a gaze-driven dual-task gaming paradigm to co-activate the active and passive attention-network competition for systematically engaging human visual-attention allocation and training. We further develop a reinforcement-learning-based human–machine gaming strategy to adjust the task parameters for improving the attention training efficiency. Consequently, we conduct an experiment study with 8 healthy participants, by jointly analyzing participants' EEG and eye-tracking data through the training process. Results show that our method can achieve improvement of brain engagement by an average of 15.6% over the widely-employed staircase strategy.**

I. Introduction

Efficient brain functional training with rehabilitation robots is an important and challenging topic in the HMI field. Existing research focus on activating the human's brain's functional behavior, such as attention, working memory, and executive function, with machine stimulation [1]–[4]. Anguera *et al.* [5] introduced an adaptive HMI approach NeuroRacer that can engage the brain activity for old adults. Wang *et al.* [6] constructed an electroencephalogram (EEG)-based HMI training paradigm to improve priority-following and modulated task-specific cortical activity in stroke patients. Chiossi *et al.* [7] built an EEG-driven adaptive VR-based HMI system that modulates visual complexity to support working memory. Lingler *et al.* [8] deployed reinforcement learning (RL) for attention management in HMI to boost performance while lowering subjective workload. However, most existing approaches are not designed under the guidance of brain function characteristics, leads to insufficient activation of target functions, cannot ensure the intensity, effectiveness and accuracy of rehabilitation training.

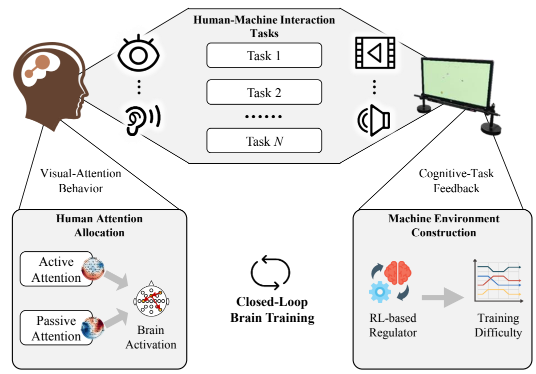


Fig. 1: Overview of the proposed visual-attention training framework. The closed-loop framework presents visible tasks that drive attention allocation and enable online adaptation.

We take the visual-attention training as an example. Visual-attention is an essential target function for brain training, which have two pathways of active and passive attention. Active attention maintenance of attentional resources initiated by internal goals, expectations, or task rules, and passive attention will capture or redirection of attentional resources by salient, novel, or abrupt external events. In HMI, the machine visual information can direct and stimulate the human's active and passive attention pathways [9]. The co-activation of these two attention pathways will drive the competition and cooperation between corresponding neural systems [10], which is the mechanism of visual-attention training. Corbetta and Shulman [11] proposed the corresponding neural basis of a dorsal–ventral dual-system. Neuroimaging also consolidates that the dorsal attention network (DAN) supports top-down selection/maintenance and the ventral attention network (VAN) reorients to salient or unexpected events [9], [12]. Beyond application settings, evidence shows a tight coupling between visual input, attentional selection, and eye movements [13]. Fei-Fei Li *et al.* [14] demonstrated that different external tasks will produce different visual-attention allocation behaviors. Therefore, the characteristic of visual-attention provides a way for systematically engaging human

This work was supported in part by the National Natural Science Foundation of China (Grant 62403104, 62473079), the Natural Science Foundation of Sichuan Province (Grant 2025ZNSFSC1518), the Fundamental Research Funds for the Central Universities (Grant ZYGX2024XJ028, ZYGX2025YGLH005), the Foundation of Nursing Key Laboratory of Sichuan Province (Grant No. HLKF2024(Z)-3).

[1]Fengjun Mu, Jingting Zhang, Zonghai Huang, Chen Chen, Chaobin Zou, Guangkui Song, Hong Cheng are with the Center for Robotics, School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, China

*Corresponding author: Jingting Zhang zhangjt@uestc.edu.cn

visual-attention training.

Adjusting the interaction and gaming behaviors between human and machine to effectively activate the brain's functional behavior is still a substantial challenge. In HMI, attention allocation varies with incoming information and is competitively reallocated when interaction tasks have changed [9], [15]. A substantial body of work has proposed training systems for medical, rehabilitation, and industrial fields [16]. Recent works [17], [18] show that gaze-based interaction techniques on eye-tracked XR systems can induce attentional selection and task performance by manipulating stimulus salience, timing, and difficulty. He *et al.* [19] demonstrated controllable changes in attention-related brain dynamics by jointly manipulating physical movement and environmental distraction, linking task context to resource reallocation at the neural level. Controlled attention-training paradigms confirmed that adapting task difficulty using gaze-sensitive metrics modulates attention allocation and improves performance [20]. Therefore, we leverage the brain function characteristics to activate the human-machine gaming, which has a remarkable capability of effectively utilizing the gaming characteristics of HMI behaviors and tasks, to effectively and precisely activate the human's active attention and passive attention for training.

In this paper, we propose a novel human-machine co-gaming interaction framework by integrating a dual-task gaming paradigm and a human–machine gaming strategy. First, we design a gaze-driven dual-task gaming paradigm to co-activate the active and passive attention-network competition for systematically engaging human visual-attention allocation and training. We further develop an RL–based human–machine gaming strategy that aims to adjust the task parameters for maintaining high brain engagement in HMI.

The key contributions of this study are threefold:

- We design a gaze-driven dual-task gaming paradigm that jointly recruits the active and passive attention pathways for competition, which systematically engaging human visual-attention allocation and training.
- We develop a human–machine gaming strategy with mechanism-guided reward shaping, to adjust the task parameters for improving the training efficiency.
- We conduct joint gaze-EEG analysis with 8 participants. Results demonstrate the co-activation of the active and passive attention-network competition, and the brain engagement improved by an average of 15.6% over the widely-employed staircase strategy.

II. Dual-Task Gaming Paradigm Design for Attention Training

Human attention control is essentially a dual-pathway system, consisting of an active attention pathway guided by internal intention and a passive attention pathway driven by external stimuli. In complex real-world environments, these two pathways are often in dynamic competition and jointly determine the allocation of attention. Our proposed dual-task gaming paradigm aims to exploit and enhance this core phenomenon under controlled experimental conditions by inducing competition for limited neural resources, thereby enhancing attentional control.

A. Mechanism of Human Visual-Attention Allocation

Human attention allows individuals to sift through this vast array of visual information, enabling them to actively focus on task-relevant contextual information to facilitate interaction, and passively switch and refocus in response to new and emergent information. This mechanism ensures efficient behavioral responses. Attention can be divided into active and passive forms, which are primarily supported by DAN and VAN respectively.

- **Active Attention from DAN:** Active attention refers to the goal-directed allocation of cognitive resources, enabling individuals to selectively focus on task-relevant information such as goals. This form of attention is supported by the DAN, a bilateral frontoparietal system primarily encompassing the intraparietal sulcus (IPS) and the frontal eye fields (FEF). The DAN is activated in situations that demand sustained attention to specific spatial locations or stimulus features.
- **Passive Attention from VAN:** The VAN mediates stimulus-driven attention allocation. It is engaged by salient or unexpected events that involuntarily capture attention. This network allows individuals to rapidly reorient their focus in response to changes in the environment.

As shown in Fig. 2, the co-occurrence of active and passive attentional mechanisms shapes the allocation of attention under the control of frontoparietal control network (FPCN), resulting in a network-level competition that serves as a driving force for controlled brain activation.

B. Dual-Task Gaming for Brain Activation

Motivated by the mechanism of human attention allocation, we introduce a gaze-driven dual-task gaming paradigm that deliberately co-engages DAN- and VAN- mediated processes to induce structured neural competition and thereby enable controllable brain activation. As shown in Fig. 3, we employ the interactive visual tasks and user monitoring to create a visually dynamic environment that simultaneously creating goal-directed and stimulus-driven attentional gaming tasks to induces structured neural competition and engages the training of attention control.

- **Goal-Directed Task for Active Attention:** For DAN activation, users are instructed to maintain visual fixation on a guidance point following a hidden trajectory. This task persistent internal attracts the participants' active attention. The motion pattern of the guidance point (e.g., velocity, acceleration) modulates DAN engagement.
- **Stimulus-Driven Task for Passive Attention:** For VAN activation, we create abrupt, high-salience elements such as blinking animations in the visual field for dynamically attract passive attention.

This dual-task design establishes a controlled regime of competition between goal-directed and stimulus-driven

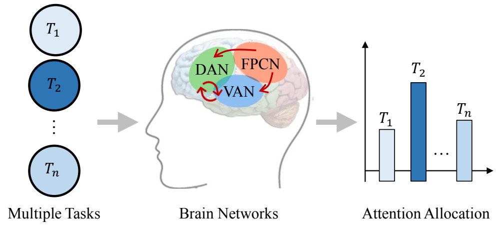


Fig. 2: Attention allocation on multiple tasks are controlled by brain networks. Multiple tasks induce competition between attention networks—goal-directed DAN and stimulus-driven VAN—coordinated by FPCN, leading to dynamic reweighting of resources toward the prioritized task.

processes, within which the balance of attention naturally evolves during interaction. The adjusting of the task difficulty and salience will sustain an appropriate level of competition and engagement, thereby fostering stable, repeatable alternation of DAN–VAN activity characteristic of effective attention training.

C. Human-Machine Gaming for Visual-Attention Training

We cast attention training as a human–machine gaming process to utilize the above principles. In this gaming process, the machine provides two visual gaming tasks that differentially recruit the human attention allocation from goal-directed and stimulus-driven processes, and parallel monitor the user's attention allocation for maintaining a controlled level of competition between the two tasks by adjusting task parameters. In this formulation, training arises from repeatedly resolving structured competition between the two processes, fostering stable, individualized practice of attentional control.

- **Gaze Tracking as Goal-Directed Task:** The user maintains gaze on a guidance point that follows a smooth, quasi-random trajectory. Task difficulty is controlled by the target kinematics P_1 and P_2. Online readouts between the goal and gaze points can quantify the sustained, top–down allocation performance.
- **Visual Salience as Stimulus-Driven Task:** While tracking continues, abrupt high-salience transients (brief flashes/motion bursts) appear at unpredictable locations and times to elicit exogenous capture. Salience and load are governed by P_3 and P_4. Behavioral measures—capture rate, saccadic latency to onset, and false alarms—index bottom–up reorienting pressure that competes with the goal-directed task.

As shown in Tab. I, we defined a set of predefined difficulties for each task, which controls key properties of the tasks in human-machine gaming. These parameters form the controllable settings for tuning the gaming environment in which DAN and VAN operate.

By adjusting these parameters based on real-time cognitive state estimates, the platform achieves a fine-grained balance between attentional load and neural plasticity, supporting sustained engagement within the user's zone of proximal

TABLE I: Dual-task human–machine gaming parameters and predefined settings to shape competition between goal-directed and stimulus-driven processes.

Task Settings		Meaning	Values
Goal-Directed Task	P_1	Max angular velocity (deg/s)	[20, 60]
	P_2	Velocity change rate (Hz)	[0.1, 0.5]
Stimulus-Driven Task	P_3	Distractor density (items)	[2, 4, 8, 16]
	P_4	Onset salience (size)	[0.5, 1, 2, 4]

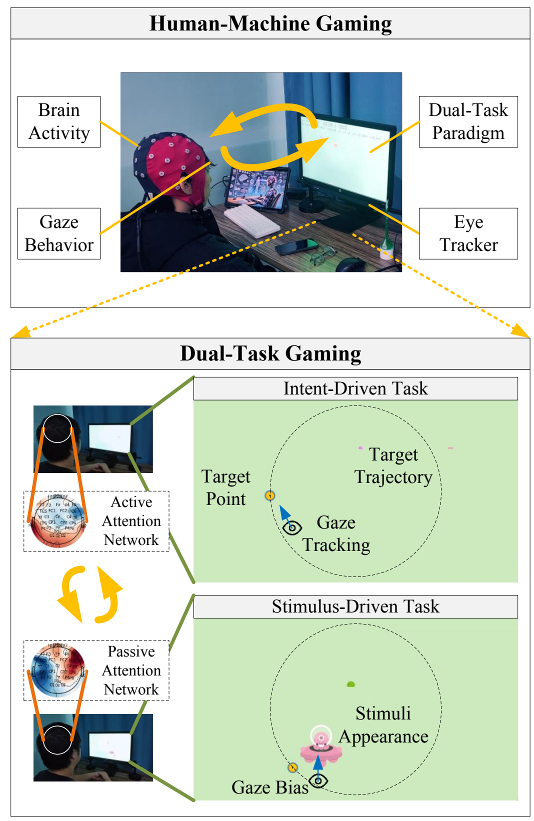


Fig. 3: Framework of human–machine gaming for attention training. A dual-task paradigm presents a goal-directed tracking task and a stimulus-driven onset task, co-engaging DAN–VAN switching. Eye-tracked gaze reveals moment-to-moment attention allocation, enabling closed-loop adaptation of task parameters and sustaining structured competition.

development. As shown in Fig. 3, in the human-machine gaming approaches, the brain activity of attention allocation status will influence the human's gaze behavior, which can be captured with an eye-tracker. Based on the human's cognitive status, the machine will provide the appropriate task settings for engaging the active and passive attention, which will activate the gaming between the corresponding brain networks for training.

In this process, the gaming between of human-machine will activate the high-level executive control system FPCN to regulate the brain activity directed toward behavioral responses, coordinates and arbitrates the competition between

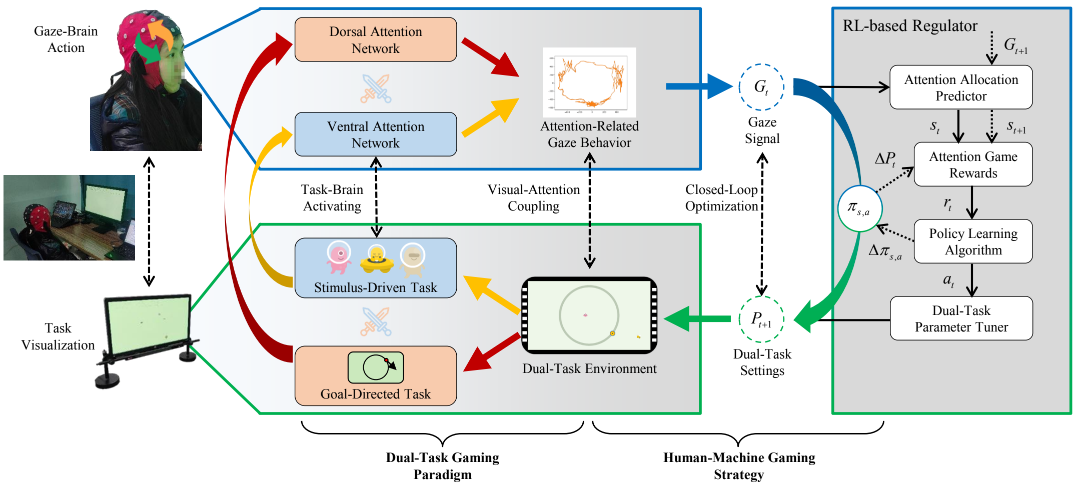


Fig. 4: **Framework Overview**. Our system is constructed with dual-task paradigm and cognitive game regulation stages. **Task–Brain Activating**: The goal-directed task competes with the stimulus-driven task, forming a controllable dual-task gaming environment that elicits resource competition between DAN and VAN under FPCN control. **Visual–Attention Coupling**: The human's attention allocation status will be reflected in gaze behavior, which can be employed as an proxy for attention allocation. **Closed-Loop Optimization**: The dual-task settings will adapt to the human-machine gaming status, for maintaining structured competition and progressively challenging training.

these parallel tasks. These task-level manipulations provide a structured mechanism for mobilizing attention-related brain networks under controlled competition. By adjusting the difficulty of this set of antagonistic tasks based on user performance, we achieve principled control of attentional dynamics. This paradigm can promote Hebbian learning-based neural plasticity, resulting in long-term benefits for users' attentional control abilities in human-machine gaming.

The components illustrated in Fig. 4 constitute a dual-task, human–machine gaming framework for attention training. By presenting concurrent goal-directed and stimulus-driven demands, the paradigm reliably co-engages DAN and VAN under FPCN supervision, yielding controlled neural competition. These dynamics are expressed in eye-movement behavior, providing measurable proxies with which the system maintains a productive level of competition. In effect, each successful suppression of distraction and restoration of focus becomes a deliberate training event, transforming moment-to-moment gaming resolution into sustained practice that strengthens attentional control.

III. Human-Machine Dual-Task Gaming Strategy with Reinforcement Learning

To reliably elicit gaming between active and passive attention and repeatedly engage the neural systems of attentional control, task parameters must be adapted online to the human–machine game state, maximizing recruitment of the target mechanisms while avoiding overload or disengagement. We learn a strategy π^* that maps the human–machine gaming state to closed-loop adjustments of the training tasks.

A. State Space from Human-Machine Interaction

Let $\mathcal{S} = \mathcal{S}_H \times \mathcal{P}$ denote the state space of human-machine gaming, where $\mathcal{S}_H = \{\mathrm{DAN}, \mathrm{VAN}\}$ is the human attention state and $\mathcal{P}$ is the discrete set of dual–task settings. At time step k, the game state is $S_k = (S_H(k), P(k)) \in \mathcal{S}$.

- **Attention Allocation State (S_H) from Human:** Exploiting the coupling between attention and gaze behavior, we infer the dominant pathway in human-machine gaming from task-specific gaze performance. Over a sliding window we compute $E = (\mu_{RT}, \sigma^2_{RT}, \mu_{TE})$, comprising mean RT-to-onset, RT variance, and mean tracking error. At step k, set $S_H(k) = 1$ (DAN-dominant) if $E \preceq \boldsymbol{\tau}$, otherwise $S_H(k) = 0$ (VAN-dominant), where $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ is subject-specific and estimated at initialization.
- **Dual-Task Settings (P) from Machine:** The game settings quantify the attentional pull each task exerts on its targeted pathway; in the human–machine game this parameterization serves as the machine-side control lever that tunes DAN–VAN competition. Following Tab. I, $P = [P_1, P_2, P_3, P_4]$ controls tracking kinematics (P_1, P_2) and distractor load/salience (P_3, P_4); each is discretized into fixed gears, yielding a finite state space.

Consequently, the joint game state is finite and discrete, spanning $\#S_H \times \#P = 2\times2\times2\times4\times4 = 128$ configurations.

B. Reward Function for Human-Machine Gaming

To guide policy exploration, we adopt a composite reward grounded in the neurocognitive goal of repeatedly engaging frontoparietal control via DAN–VAN gaming resolution.

At time step k the overall reward is

$$R(k) = \sum_{i=1}^{4} \beta_i \, r_i(k), \tag{1}$$

where β_i are weights reflecting the relative importance of corresponding reward items as follows:

a) r_1 for Dual-Task Gaming: This term rewards parameter updates that purposefully shape competition between the goal-directed and stimulus-driven tasks in accordance with the current attentional state. For example, when DAN dominant, actions aim to boost VAN activation via the stimulus-driven pathway; when VAN dominant, the policy should aim to re-engage DAN via the goal-directed pathway. Thus, r_1 will maintains a dual-task gaming rather than simply using easy tasks to maximize behavioral performance:

$$r_1(k) = (\Delta\bar{P}(k) \oplus S_H) + \alpha_1\Sigma(P(k)), \tag{2}$$

where $\Delta\bar{P}(k)$ is the parameter change, $\oplus$ denotes the exclusive-or, and $\Sigma(\cdot)$ measures normalized challenge.

b) r_2 for Human-Machine Gaming: This term values how the closed loop stabilizes focus and times interventions to promote effective recovery after distraction. It jointly rewards sustained goal-directed engagement and successful re-engagement following a switch:

$$L(k) = \max\{n \in \mathbb{N} \mid S_H(k) = \ldots = S_H(k-n+1)\}, \tag{3}$$

$$r_2(k) = \arctan\left(\alpha_2(2S_H(k) - 1) + b\Delta S_H(k) \cdot L(k)\right) \tag{4}$$

where $L(k)$ is the dwell length of the current state and $\Delta S_H(k)$ indicates state transitions.

c) r_3 for safety: This term imposes a penalty when task parameters exceed predefined safety bounds in Table I, preventing cognitive overload.

d) r_4 for anti-stagnation: This term penalizes overly static training conditions, as prolonged periods under a single setting may encourage habitual responding rather than engaging attentional control.

The composite reward steers the policy toward neurally targeted and behaviorally interpretable control rather than mere score maximization. By combining the regulation of dual-task gaming (r_1), human-machine gaming (r_2), safety (r_3), and anti-stagnation (r_4), it aligns learning with our central objective: strengthening frontoparietal control via context-sensitive co-engagement of the DAN–VAN system.

C. Action and Policy Learning

We cast parameter adaptation as a sequential decision problem. At each decision step the agent observes the current game state and selects a small adjustment to the task configuration. The design goals are interpretability, smoothness, and safety so that competition between goal-directed and stimulus-driven processes is sustained without overloading the participant.

a) Action Space: At each decision step the agent modifies the dual-task settings in a minimal manner to ensure smooth human–machine gaming. It may increase or decrease a single task setting, or leave the settings unchanged. Therefore, the action space contains $2 \times 4 + 1 = 9$ discrete actions.

b) Policy Learning: With the state defined in Sec. III-A, the attention training process is modeled as a Markov decision process. A transition results from applying a feasible adjustment, observing the subsequent attention allocation and reward, and advancing to the next step. We learn an approximately optimal policy over the finite state and action spaces using tabular Q-learning with safety masking and ϵ-greedy exploration. Policies are trained online for each participant. Exploration is reduced over time, and a small minimum dwell window is enforced before reversing a recent change to improve smoothness. All hyperparameters and schedules are reported in the appendix.

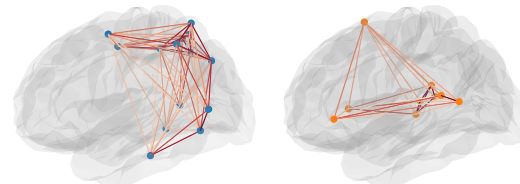
IV. Experiment

To evaluate the efficacy of the proposed human–machine dual-task visual-attention training paradigm, we conducted a controlled study comprising neurophysiological and behavioral assessments, with EEG recorded throughout.

A. Participants and Data Acquisition

Eight healthy participants (6 males, 2 females; aged 21-28 years) voluntarily participated in this study. All subjects had normal or corrected-to-normal vision and reported no history of neurological disorders that could affect experimental outcomes. Written informed consent was obtained from all participants prior to the experiment, and the study protocol was approved by the independent ethics committee of University of Electronic Science and Technology of China (No. 34083).

To monitor the human activity in human-machine gaming, EEG signals were recorded using a 32-channel eego mylab system from ANT Neuro. Signals were sampled at 1000 Hz, notch-filtered at 50 Hz to remove power-line interference. Eye movements were recorded with a Tobii Pro Fusion eye tracker mounted below a monitor positioned 70 cm from the participant's eyes, at a sampling rate of 120 Hz. The tracker was calibrated using a standard 9-point procedure. The proposed dual-task gaming paradigm is implemented with PyGame framework with a fixed frame rate of 50 Hz, and human-machine dual-task gaming strategy is explored on single participant for 40 minutes.



(a) DAN (b) VAN

Fig. 5: Canonical attention networks. [21] (a) Dorsal attention network (DAN) and (b) ventral attention network (VAN) rendered on three views of the cortex.

B. Metrics for Brain Activity and Visual-Attention Behavior

In our experiment, we extracted interpretable features of participants' eye movement behavior and neural activity and calculated the mutual information metric:

- **Spectral Activation Ratio (SAR):** Cortical activation can be estimated using the β/α power ratio. This ratio increases when the brain is activated and decreases during brain inactivation, makes it a proxy of cortical engagement.
- **Directional Alignment Index (DAI):** This metric characterizes the visual-attention allocation by analyzing the directional consistency between gaze movement and the trajectory of the visual tracking target. At each time step t, the cosine similarity between the gaze vector $v_{\text{gaze}}(t)$ and the local tangential direction of the target $v_{\text{target}}(t)$ is computed as:

$$\text{DAI}(t) = \frac{v_{\text{gaze}}(t) \cdot v_{\text{target}}(t)}{\|v_{\text{gaze}}(t)\| \cdot \|v_{\text{target}}(t)\|},$$

  where $v_{\text{gaze}}(t)$ and $v_{\text{target}}(t)$ is the direction vector of gaze and target, respectively. Higher DAI indicates better performance on the gaze tracking task.
- **Mutual Information (MI):** To capture neural–behavioral coupling, MI was calculated between network dominance (DAN vs. VAN activation in Fig. 5) and gaze allocation (target vs. distractor), indicating the dependency between internal neural states and external attention behavior.

C. Gaze-EEG Coupling Analysis in Visual-Brain Training

To validate the effectiveness of the proposed gaze-driven dual-task gaming paradigm for brain training, we analyzed the coupling between brain activity and gaze behavior.

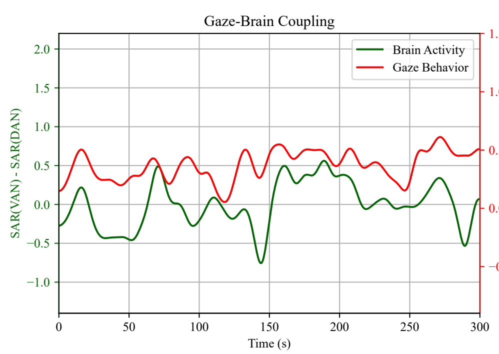


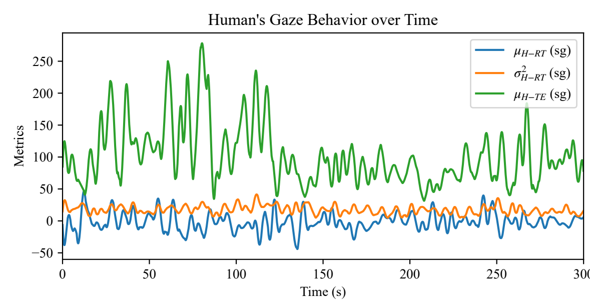
Fig. 6: Temporal coupling between the visual behavior and attention activity in DAN and VAN indicating that eye–movement deviation tracks the neural attention balance.

As shown in Fig. 6, DAI induces the human's performance on the gaze tracking task, and the comparison between DAN and VAN indicates the brain activity of attention allocation. To analyze the vision-attention coupling, we analyzed the two time series. Results exhibit a reliable association: Pearson correlation $r = 0.515$ with $p < 0.001$ and Spearman correlation $r = 0.552$ with $p < 0.001$, indicating both linear and monotonic coupling. Information–theoretic analysis further supports this link. The mutual information (MI) between the gaze feature DAI and the brain competition is 0.850, and pathway-specific dependencies are consistent with our design, with $MI(\text{DAN},\ \text{DAI}) = 0.656$ and $MI(\text{VAN},\ \text{1-DAI}) = 0.832$.

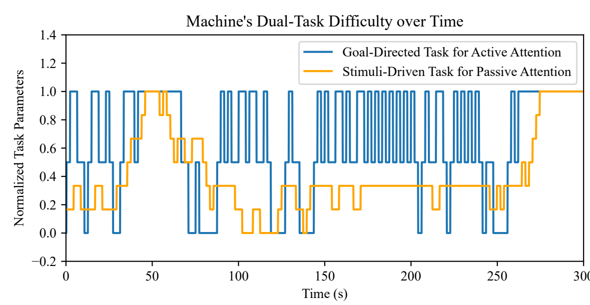
These results show that the gaze feature can reliably tracks the evolving balance between goal-directed and stimulus-driven processes, which justifies using gaze as an online proxy for attention allocation and validates the closed-loop adaptation mechanism employed by our dual-task controller.

D. Active- and Passive-Attention Activation in Dual-Task Gaming Paradigm

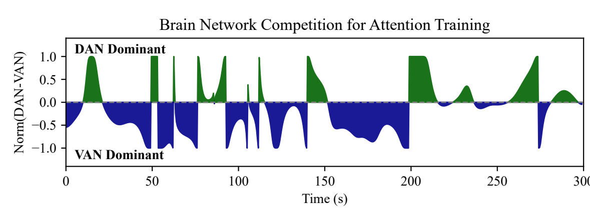
We calculated the time-varying dominance of the DAN and VAN using EEG data. The spectral activation ratio (SAR) in the beta band (13–30 Hz) was computed for electrode clusters corresponding to DAN (F3, Fz, F4, P3, Pz, P4) and VAN (T7, T8, CP5, CP6). Fig. 7 illustrates the dynamic interplay between these networks during dual-task gaming, under the control of the proposed human-machine dual-task strategy.



(a) Realtime Gaming Metrics from Gaze Observation. RT error is measured in frames, while TE is measured in pixels.



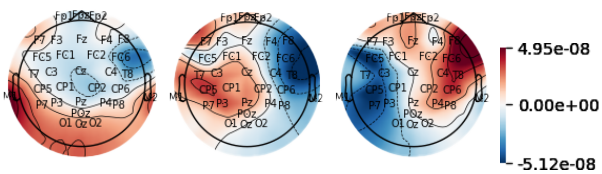
(b) Dynamic Task Adaptation to Maintain Attention Engagement.



(c) DAN-VAN Network Competition under Dual-Task Gaming.

Fig. 7: Closed-loop behavior–task–brain dynamics in attention training.

During interaction, the participant's allocation of attention evolves over time, which is reflected by dynamic changes in gaze-based behavioral metrics in Fig. 7a (nonstationary trends in tracking error and reaction–time statistics rather than a fixed level). In response to these behavioral fluctuations, the controller adapts the dual-task settings in Fig. 7b through small, single-gear adjustments to goal-directed difficulty and stimulus-driven salience, thereby reshaping the visual input and the competitive pressure presented to the user. As expected, this induces corresponding shifts in network dominance in Fig. 7c, where the competition index alternates between DAN- and VAN-dominant regimes with bounded dwell lengths, indicating sustained co-engagement rather than trivialization or instability. This behavior–task–brain cascade motivates a quantitative analysis of the underlying vision–attention coupling, which we present in Sec. IV-D.



(a) t=150s (b) t=151s (c) t=152s

Fig. 8: Temporal-spatial evolution of EEG topographies during a DAN-to-VAN transition in dual-task gaming.

On the temporal-spatial aspect, we further analyzed the temporal-spatial brain activity within the dual–task gaming. As indicated in Fig. 7c, at 151s the attention dominance shifts from DAN to VAN.

Fig. 8 support this transition in the spatial domain:

- **Goal-Directed Stage**: Fig. 8a shows activity concentrated over posterior parietal–occipital and midline sites with comparatively weaker lateral temporal responses, consistent with DAN engagement.
- **Transition Stage**: Fig. 8b shows attenuation over posterior parietal regions and the emergence of lateral temporal–parietal clusters, marking the transition.
- **Stimulus-Driven Stage**: Fig. 8c is dominated by ventral frontotemporal and temporoparietal activity with reduced dorsal parietal contribution, aligning with VAN-dominant reorienting driven by the salience stimuli.

The temporal-spatial evolution demonstrates that the dual-task controller drives rapid and repeatable shifts between DAN and VAN, providing direct neurophysiological evidence that the paradigm modulates attention networks as designed.

E. Training Engagement Analysis under Human-Machine Gaming Strategy

To investigate the neurophysiological underpinnings of our proposed dual-task paradigm, we analyzed EEG data from eight healthy subjects. Our primary objective was to determine whether the reinforcement learning-based strategy (HRL) could induce a higher level of cortical arousal compared to the staircase algorithm (SA) which is widely employed by other training approach [5]. We operationalized cortical arousal and attention engagement with a well-established neural marker: the power ratio of beta frequency band (13–30 Hz) to alpha frequency band (8–12 Hz) over a set of prefrontal and parietal channels.

As visually summarized in Fig. 9, the HRL strategy consistently elicited a higher attention engagement level across the majority of subjects. The individual time-series plots further reveal the dynamic nature of attention reallocation induced by the dual-task paradigm. The quantitative results also confirm this trend. The mean attention level for the HRL condition ($M = 0.4137$, $SD = 0.1499$) was substantially higher than that for the SA condition ($M = 0.3579$, $SD = 0.1268$). This corresponds to a 15.6% relative increase in brain engagement under the HRL strategy compared to the SA baseline.

To rigorously assess the significance of this difference, we conducted a paired-samples t-test. Prior to the test, a Shapiro-Wilk test was performed on the distribution of the paired differences, which confirmed its normality ($W = 0.8988$, $p = 0.2818$), thus validating the use of a parametric test. The t-test revealed a statistically significant difference between the two conditions, $t = 3.0295$, $p = 0.019 < 0.05$, indicating that the HRL strategy led to a significantly greater level of cortical arousal. Furthermore, the calculated effect size was very large (Cohen's $d = 1.0711$), which underscores the practical and robust impact of the HRL strategy on attentional engagement.

These neurophysiological findings provide compelling evidence that the proposed human-machine dual-task strategy successfully induces attention reallocation. They robustly demonstrate that the adaptive, reinforcement learning-based HRL strategy is superior to a static staircase strategy in enhancing and sustaining cortical engagement during training. This suggests that such an adaptive approach holds significant promise for developing more effective attention training and neurorehabilitation protocols.

V. Conclusion

This paper presented a visual–attention training paradigm by realizing a closed-loop HMI co-gaming system. Based on the mechanism of human's visual-attention generation, we propose a dual-task gaming paradigm that employ goal-directed and stimulus-driven tasks to co-activate the active and passive attention-network competition for systematically engaging human visual-attention allocation and training. In this paradigm, we employ a human-machine dual-task gaming strategy that employ gaze signal from eye-tracker to determine an online proxy for attention allocation. We further propose a human–machine gaming strategy to adjust the task parameters for improving the attention training efficiency. To demonstrate the effectiveness of our system, we conducted a comprehensive study with 8 participants. Results indicate that the framework reliably induces and regulates attention allocation and sustains higher brain engagement than a widely-employed staircase baseline, suggesting a principled route to individualized, mechanism-driven visual-attention training. In the future, we plan to validate on larger participant cohorts and expand the study to probabilistic

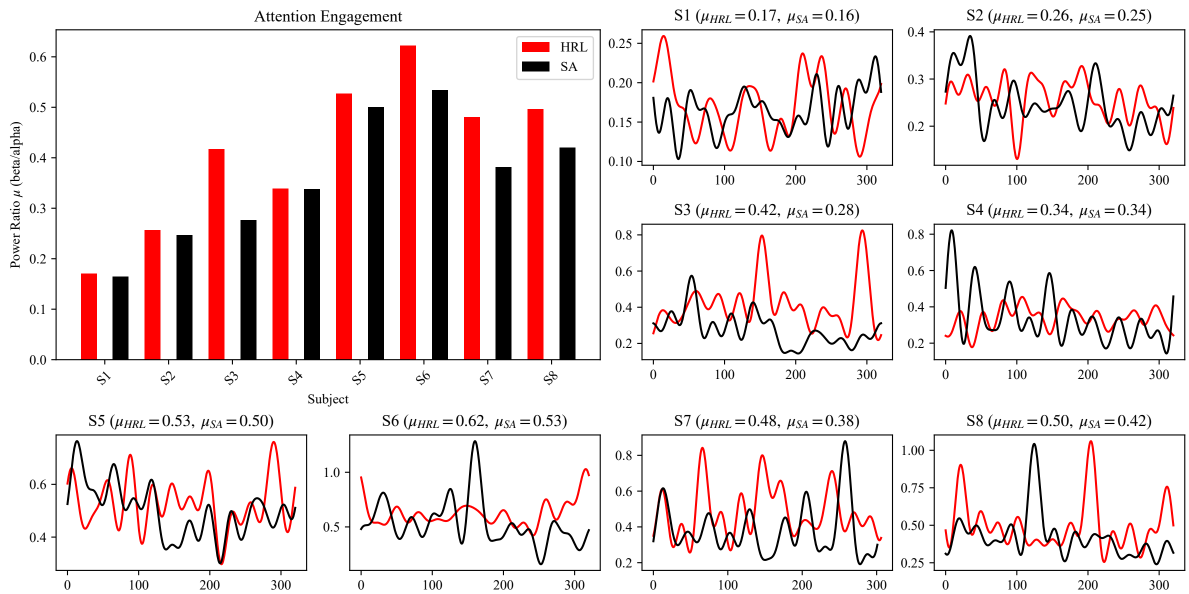


Fig. 9: Attention engagement of β/α over fronto-parietal electrodes (F3, Fz, F4, P3, Pz, P4) between the proposed Human-Machine Dual-Task Gaming Strategy (HRL) and Staircase Algorithm (SA). Across-subject summary bars and per-subject trajectories with average attention engagement level μ_{HRL} and μ_{SA} are annotated.

continuous state estimation with multi-modal sensing, finer-grained control policies to strengthen robustness, safety, and translational impact.

References

[1] J. A. Anguera, M. A. Rowe, J. J. Volponi, M. Elkurdi, B. Jurigova, A. J. Simon, R. Anguera-Singla, C. L. Gallen, A. Gazzaley, and E. J. Marco, "Enhancing attention in children using an integrated cognitive-physical videogame: A pilot study," *npj Digital Medicine*, vol. 6, no. 1, Art. 65, Apr. 2023.

[2] K. Shi, R. Huang, F. Mu, Z. Peng, K. Huang, Y. Qin, X. Yang, and H. Cheng, "A Novel Multimodal Human-Exoskeleton Interface Based on EEG and sEMG Activity for Rehabilitation Training," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2022, pp. 8076–8082.

[3] C. Xu, C. Zhang, Y. Zhou, Z. Wang, P. Lu, and B. He, "Trust Recognition in Human-Robot Cooperation Using EEG," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2024, pp. 7827–7833.

[4] Z. Yu, V. Nalam, A. Alili, and H. H. Huang, "A Wearable Robotic Rehabilitation System for Neuro-Rehabilitation Aimed at Enhancing Mediolateral Balance," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2023, pp. 1–6.

[5] J. A. Anguera, J. Boccanfuso, J. Rintoul, O. Claflin, J. Naranjo, F. A. F. Camilleri, E. L. Gazzaley, S. M. Chiong, *et al.*, "Video game training enhances cognitive control in older adults," *Nature*, vol. 501, no. 7465, pp. 97–101, 2013.

[6] J. Wang, W. Wang, and Z.-G. Hou, "EEG-Based Focus of Attention Tracking and Regulation During Dual-Task Training for Neural Rehabilitation of Stroke Patients," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 3, pp. 920–930, Mar. 2023.

[7] F. Chiossi, C. Ou, C. Gerhardt, F. Putze, and S. Mayer, "Designing and evaluating an adaptive virtual reality system using EEG frequencies to balance internal and external attention states," *International Journal of Human–Computer Studies*, 2025, Art. no. 103433.

[8] A. Lingler, P. Langerak, H.-K. Choi, S. Mayer, and W. Stuerzlinger, "Supporting task switching with reinforcement learning," in *Proc. CHI '24: ACM Conference on Human Factors in Computing Systems*.

[9] A. Tosoni, P. Capotosto, A. Baldassarre, S. Spadone, and C. Sestieri, "Neuroimaging evidence supporting a dual-network architecture for the control of visuospatial attention in the human brain: a mini review," *Frontiers in Human Neuroscience*, vol. 17, 1250096, 2023.

[10] J. D. Power, A. L. Cohen, S. M. Nelson, *et al.*, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.

[11] M. Corbetta and G. L. Shulman, "Control of goal-directed and stimulus-driven attention in the brain," *Nature Reviews Neuroscience*, vol. 3, no. 3, pp. 201–215, 2002.

[12] J. Zhao, J. Wang, C. Huang, and P. Liang, "Involvement of the dorsal and ventral attention networks in visual attention span," *Human Brain Mapping*, vol. 43, no. 6, pp. 1941–1954, 2022.

[13] R. N. Denison, M. Carrasco, and D. J. Heeger, "A dynamic normalization model of temporal attention," *Nature Human Behaviour*, vol. 5, pp. 1674–1685, 2021.

[14] F. F. Li, R. VanRullen, C. Koch, and P. Perona, "Rapid natural scene categorization in the near absence of attention," *Proc. Natl. Acad. Sci. USA*, vol. 99, no. 14, pp. 9596–9601, 2002.

[15] J. Theeuwes, "Attentional capture and control," *Annual Review of Psychology*, 2025.

[16] I. M. Butaslac, Y. Fujimoto, T. Sawabe, M. Kanbara, and H. Kato, "Systematic Review of Augmented Reality Training Systems," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 12, pp. 5062–5082, Dec. 2023.

[17] C. Ebner, A. Plopski, D. Schmalstieg, and D. Kalkofen, "Gaze-Contingent Layered Optical See-Through Displays with a Confidence-Driven View Volume," *IEEE Trans. Vis. Comput. Graph.*, vol. 30, no. 11, pp. 7203–7213, 2024.

[18] S. Chen, B. Duinkharjav, X. Sun, L.-Y. Wei, S. Petrangeli, J. Echevarria, C. Silva, and Q. Sun, "Instant Reality: Gaze-Contingent Perceptual Optimization for 3D Virtual Reality Streaming," *IEEE Trans. Vis. Comput. Graph.*, vol. 28, no. 5, pp. 2157–2167, 2022.

[19] C. He, Y.-Y. Chen, C.-R. Phang, I.-P. Chen, S.-C. Tzou, T.-P. Jung, and L.-W. Ko, "Exploring Embodied Cognition and Brain Dynamics Under Multi-Tasks Target Detection in Immersive Projector-Based Augmented Reality (IPAR) Scenarios," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 3476–3485, 2024.

[20] D. Rane, P. Sharma, M. Singh, and U. Lahiri, "Virtual Reality Based Gaze-Sensitive Aiming Task Platform: Role of Attention Allocation in Task Performance for Individuals With Autism and Typically Developing Individuals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1492–1501, 2023.

[21] B. A. Seitzman, C. Gratton, T. O. Laumann, E. M. Gordon, B. Adeyemo, A. Dworetsky, B. T. Kraus, A. W. Gilmore, J. J. Berg, M. Ortega, A. Nguyen, D. J. Greene, K. B. McDermott, S. M. Nelson, C. N. Lessov-Schlaggar, B. L. Schlaggar, N. U. F. Dosenbach, and S. E. Petersen, "Trait-like variants in human functional brain networks," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 116, no. 45, pp. 22851–22861, 2019.