International Journal of Computer Applications (0975 – 8887)
Volume 187 – No.29, August 2025

VALF: A CLA-based Reinforcement Learning Framework for Vietnamese Language Learning in Dyslexic Children

Vũ Trần Quang Minh
UG Scholar, Computing Department,
University of Greenwich
alliance with FPT University,
Ho Chi Minh City, Vietnam

Sharmila Mathivanan
IT Lecturer, Faculty of Computing Department,
University of Greenwich
alliance with FPT University,
Ho Chi Minh City, Vietnam

ABSTRACT
Dyslexic children faces many persistent challenges in learning Tonal languages such as Vietnamese, Chinese, Koeran where visual, phonetic, and orthographic components interact in complex ways. This paper introduces a new approach VALF (Vietnamese Adaptive Learning Framework), a novel educational AI prototype that uses Cellular Learning Automata (CLA) to generate adaptive multimedia content for Vietnamese language instruction. VALF integrates a reinforcement learning-driven virtual pen for character rendering, tone marker placement, and audio pronunciation to enhance learners' visual, auditory, and motor integration. Designed for primary school children with dyslexia, the system simulates personalized learning through games, quizzes, and pronunciation feedback. The proposed VALF model can also able to produce long texts and sentences using Vietnamese Word Generation algorithm. Main objectives of VALF's efficacy is to promote literacy and self-confidence in dyslexic learners. Future work includes implementing VALF in gaming platform to support and integrate deep learning for natural speech synthesis and feedback.

General Terms
Artificial Intelligence, Cellular Learning Systems, Adaptive Learning, Reinforcement Learning, Language Learning.

Keywords
Vietnamese language learning, Educational AI, Reinforcement learning, Cellular Learning Automata (CLA), Dyslexia, Computer-assisted language learning (CALL), Adaptive learning systems, Child education technology, Multimodal learning, Tone-based script modelling.

1. INTRODUCTION
Language learning poses distinct challenges to children diagnosed with dyslexia, a neurodevelopmental disorder characterized by persistent difficulties in accurate and fluent word recognition, poor spelling, and decoding abilities. These challenges are amplified when learning tonal languages such as Vietnamese, where the integration of phonemic, orthographic, and tonal elements significantly impacts the readability and comprehensibility of written and spoken forms. Recent studies have underscored that adaptive digital learning systems driven by artificial intelligence (AI) can significantly enhance language acquisition among learners with special educational needs by providing personalized, responsive instructional environments.

In response to this need, computer-assisted language learning (CALL) frameworks, underpinned by various machine learning methods, have seen increasing adoption. Among these, reinforcement learning (RL) methodologies stand out due to their ability to iteratively adapt content and learning interactions based on learner feedback, without dependence on extensive labelled datasets typically required by supervised learning techniques. One notable example is the Reinforcement Adaptive Learning Framework (RALF) introduced by Minoofam et al. [1], which utilizes Cellular Learning Automata (CLA), a form of reinforcement learning combined with cellular automata, for Persian language instruction tailored specifically to dyslexic children. Their study demonstrated that reinforcement learning frameworks significantly improved reading literacy and engagement among learners with dyslexia compared to conventional teaching methods, highlighting the promise of CLA-based adaptive learning systems.

However, despite this promising development, few studies have explored the application of such adaptive reinforcement learning models to tonal languages—such as Vietnamese—which present unique phonological and orthographic complexities due to their tone-dependent meaning structure. Vietnamese employs six distinct tonal markers, and slight variations in tone or orthographic structure can drastically alter meaning, adding complexity that substantially impacts language acquisition for learners with dyslexia. Thus, there is an acute need for specialized adaptive learning environments capable of effectively addressing these unique linguistic demands.

To bridge this gap, the current study proposes the Vietnamese Adaptive Learning Framework (VALF), a novel CLA-based reinforcement learning model tailored explicitly to teaching Vietnamese to dyslexic primary school learners. VALF systematically extends the pioneering approach of RALF by incorporating the complexities of Vietnamese tones, character structure, and pronunciation. The system generates adaptive educational content dynamically, employing a virtual pen to illustrate precise character formation sequences, audio-visual feedback for tonal accuracy, and interactive games that reinforce the integrative skills of reading, writing, and speaking.

This paper aims to evaluate the effectiveness of VALF through a simulated case study, analysing its potential in terms of learner engagement, skill acquisition, and instructional adaptability. By drawing on related literature from the domain of reinforcement learning in education [2], multimodal instructional technology [3], and adaptive systems for learners with special educational needs [4], we situate our work within a robust theoretical and practical context.

The remainder of this paper is organized as follows: Section II reviews relevant literature on reinforcement learning applications in CALL and dyslexia-focused adaptive systems. Section III presents foundational concepts, including characteristics of the Vietnamese language, particularly tonal and orthographic structures, alongside the theoretical underpinnings of Cellular Learning Automata. Section IV elaborates on the VALF system's architecture and adaptive content generation processes. Section V presents a simulated evaluation of VALF's educational efficacy. Section VI discusses implications, challenges, and limitations, and Section VII provides concluding remarks and directions for future research.

2. LITERATURE REVIEW

In recent years, computer-assisted language learning (CALL) has become a key area of educational technology research, driven by the widespread integration of machine learning (ML) and artificial intelligence (AI) methods. These technologies significantly enhance learner engagement, personalize learning experiences, and improve educational outcomes. Particularly, reinforcement learning (RL), a subset of machine learning, has demonstrated exceptional potential for creating adaptive, personalized learning environments without the need for extensive labeled datasets [2]. Reinforcement learning's capacity to dynamically adapt based on learner responses makes it particularly suitable for applications in special education contexts, including dyslexia intervention.

2.1 Reinforcement Learning for Dyslexia Intervention

The recent work by Minoofam et al. [1], which introduced the Reinforcement Adaptive Learning Framework (RALF), is particularly significant. RALF utilized Cellular Learning Automata (CLA), a specialized form of RL that integrates cellular automata with learning automata to generate adaptive Persian-language content specifically for dyslexic learners. This system adapted character generation, word formation, and sentence composition, significantly enhancing literacy and learner engagement compared to traditional classroom methods. Results indicated a near 27% improvement in learning performance among dyslexic children, demonstrating the substantial promise of CLA for targeted educational interventions.

RALF's success was partly due to its innovative integration of visual, auditory, and interactive gaming elements, an approach supported by findings in educational multimedia research [3], [5]. The framework provided direct, immediate feedback and dynamic adjustment of difficulty, which proved particularly beneficial for dyslexic learners who commonly struggle with conventional educational techniques [4].

2.2 Computer-Assisted Tonal Language Learning

However, tonal languages such as Vietnamese present additional complexity for learners, especially dyslexic children. Vietnamese orthography involves six distinct tones, with subtle pronunciation variations significantly changing meaning. Traditional CALL systems rarely address these specific phonetic and orthographic challenges comprehensively.

Prior studies into CALL for tonal languages often focus primarily on Mandarin Chinese. For instance, Peng et al. [6] developed a 3D virtual talking head system for Mandarin learners, significantly improving pronunciation through visual-auditory integration. Although highly successful, their approach relied on supervised learning, requiring extensive training datasets—something less feasible for resource-constrained contexts or when developing systems for learners with special educational needs.

2.3 Policy Gradient Methods

Policy Gradient (PG) methods represent a fundamental class of reinforcement learning algorithms that aim to directly optimize the policy function, which maps states to action probabilities. Rather than estimating the value of each state-action pair as in value-based methods (e.g., Q-learning), PG methods learn a parameterized policy that maximizes expected cumulative reward by ascending the gradient of the performance objective with respect to policy parameters.

Mathematically, the policy is typically represented as $\pi_\theta(a \mid s)$, where θ denotes the parameters of the policy network. The optimization objective is defined as $J(\theta) = E[R]$, where R is the return obtained from following the policy $\pi\theta$. Using the REINFORCE algorithm [7], the policy gradient can be estimated as:

$$\nabla_\theta J(\theta) \approx E[\nabla_\theta \log \pi_\theta (a \mid s) \cdot R] \qquad (1)$$

The procedural steps of this method are illustrated in Figure 1. The diagram shows how the agent gathers training data by interacting with the environment during an episode. At the end of each episode, the observed rewards are used to compute the return, which is then utilized to compute the loss and update the policy network. This cycle continues for multiple epochs to refine the agent's behavior.

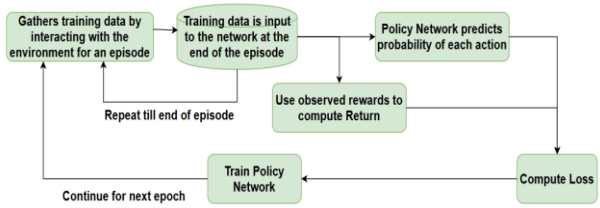

Figure 1. Workflow of policy gradient reinforcement learning

The agent gathers training data by interacting with the environment. After each episode, returns are computed using observed rewards. The policy network then computes the loss and updates its parameters to improve future performance.

This approach is particularly beneficial in continuous or high-dimensional action spaces where value-based methods may struggle. Policy gradient methods are widely applied in robotics, natural language generation, and complex decision-making environments due to their flexibility and capacity to handle stochastic policies.

However, in the context of adaptive educational systems such as VALF, which targets Vietnamese language learning for dyslexic children, the limitations of standard policy gradient methods become apparent. First, PG methods typically suffer from high variance in gradient estimates, leading to unstable learning. Second, they often require substantial amounts of interaction data, which is impractical in educational environments where learner time and cognitive load are constrained. Third, the lack of interpretability in neural policy representations poses challenges for educational stakeholders who require transparency in system behaviour.

Given these constraints, VALF employs Cellular Learning Automata (CLA), a lightweight, interpretable reinforcement learning model that aligns well with the symbolic and spatial characteristics of Vietnamese script and tonal structure. Unlike policy gradient methods, CLA updates policies at the cell level through discrete actions and localized feedback, enabling fine-grained control over character rendering and tone placement while maintaining computational efficiency and explainability. This makes CLA a more pedagogically appropriate choice for developing personalized, accessible learning tools for dyslexic learners.

2.4 Q-Learning

Q-learning is one of the most widely used value-based reinforcement learning algorithms, originally proposed by Watkins in 1989 [8]. It is an off-policy temporal difference (TD) learning method that enables an agent to learn the optimal action-selection policy by estimating the value of state-action pairs, known as Q-values. The core idea of Q-learning is to iteratively update the Q-values using the Bellman equation:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q\,(s',a') - Q(s,a)\right] \quad (2)$$

where:

- s and a are the current state and action,
- r is the reward received,
- s' is the next state,
- γ is the discount factor,
- α is the learning rate.

By repeatedly applying this update rule, the agent converges to the optimal policy, even without prior knowledge of the environment model.

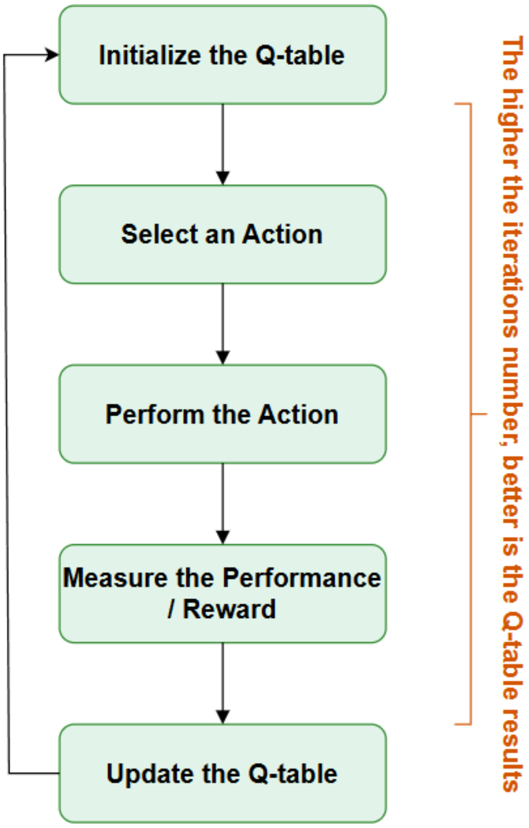

Figure 2. Working process of the Q-learning algorithm

The agent iteratively updates the Q-table based on rewards received after performing actions in the environment. Over multiple iterations, the policy converges toward optimal behaviours.

Q-learning has been applied in numerous domains such as robotics, game playing, recommendation systems, and control tasks [10], [11]. Its success is largely attributed to its simplicity, model-free nature, and guaranteed convergence under certain conditions [9].

However, Q-learning also presents several challenges when applied to complex or high-dimensional environments. It requires maintaining a Q-table whose size grows exponentially with the number of states and actions, making it impractical for large state spaces. This limitation has led to extensions such as Deep Q-Networks (DQN), which use neural networks to approximate Q-values [11].

In the context of educational applications like VALF, Q-learning's tabular formulation is not ideally suited for tasks involving structured symbolic representations such as Vietnamese orthography and tone placement. Additionally, Q-learning assumes discrete action and state spaces, whereas educational content generation may involve more complex dependencies between visual, phonetic, and auditory elements.

Given these constraints, VALF instead employs Cellular Learning Automata (CLA), which offers localized, explainable decision-making and better aligns with the spatial and symbolic structure of language learning tasks.

2.5 Deep Q-Networks (DQN)

Deep Q-Networks (DQN) represent a major advancement in reinforcement learning by combining Q-learning with deep neural networks to handle environments with high-dimensional or continuous state spaces. The concept was popularized by the seminal work of Mnih et al. [11], where a convolutional neural network (CNN) was used to approximate Q-values for playing Atari 2600 games directly from raw pixel input.

In traditional Q-learning, a Q-table is used to store Q-values for all state-action pairs. However, this becomes infeasible in environments with large or continuous state spaces. DQN addresses this limitation by using a neural network parameterized by θ to approximate the Q-function: $Q(s,a;\theta)$. The network takes a state as input and outputs Q-values for all possible actions.

Key innovations introduced in DQN include:

- **Experience Replay:** A buffer stores transitions (s, a, r, s'), and the network is trained on random batches from this buffer, breaking correlations between sequential observations (Lin, 1992).
- **Target Network:** A separate network is used to compute target Q-values, and its parameters are updated periodically, stabilizing training (Mnih et al., 2015).

The training objective in DQN minimizes the difference between the predicted Q-value and the target Q-value using the loss function:

$$L(\theta) = E\left[\left(r + \gamma \max_{a'} Q\,(s',a';\theta^{-}) - Q(s,a;\theta)\right)^2\right] \quad (3)$$

where θº denotes the parameters of the target network.

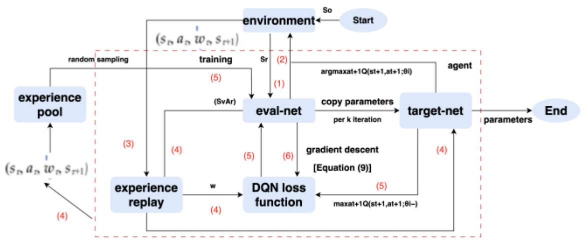

Figure 3. Deep Q-Network (DQN) architecture and training process

The agent interacts with the environment, storing transitions in the experience pool, which are sampled to train the eval network. A target network is periodically updated to stabilize training.

DQN and its extensions—including Double DQN [13], Dueling DQN [14], and Prioritized Experience Replay [15]—have achieved state-of-the-art performance in a wide range of tasks such as video game playing, robotic control, and autonomous navigation.

Despite its success, DQN also faces challenges such as overestimation bias, sample inefficiency, and slow convergence. In educational applications like VALF, DQN's complexity and lack of interpretability pose limitations. Since VALF emphasizes transparent, symbolic, and spatial learning for dyslexic students, using a black-box model like DQN may obscure the learning logic and hinder pedagogical trust.

Instead, VALF leverages Cellular Learning Automata (CLA), which allows fine-grained, cell-level updates and visual traceability in educational contexts.

2.6 Actor-Critic Methods

Actor-Critic methods are a class of reinforcement learning algorithms that combine the benefits of both value-based and policy-based approaches. Instead of learning a value function or a policy in isolation, Actor-Critic architectures learn both simultaneously: the actor updates the policy function to select actions, and the critic evaluates the actions taken by estimating the value function [16]. The actor is responsible for learning the policy $\pi(a|s;\theta)$, which maps states to action probabilities, while the critic learns a value function $V(s;w)$ or an action-value function $Q(s,a;w)$ that provides feedback to improve the actor's policy. The two components are trained jointly using policy gradients and temporal difference (TD) errors. The policy is updated by ascending the gradient:

$$\nabla_\theta J(\theta) = E[\nabla_\theta \log \pi(a \mid s;\theta) \cdot \delta] \qquad (4)$$

where $\delta = r + \gamma V(s';w) - V(s;w)$ is the TD error estimated by the critic. Actor-Critic algorithms address some of the limitations of pure policy gradient methods (e.g., high variance) and pure value-based methods (e.g., poor exploration). Variants such as Advantage Actor-Critic (A2C), Asynchronous Advantage Actor-Critic (A3C) [17], Deep Deterministic Policy Gradient (DDPG) [18], and Soft Actor-Critic (SAC) [19] have been successfully applied to continuous control tasks, robotics, and video game AI.

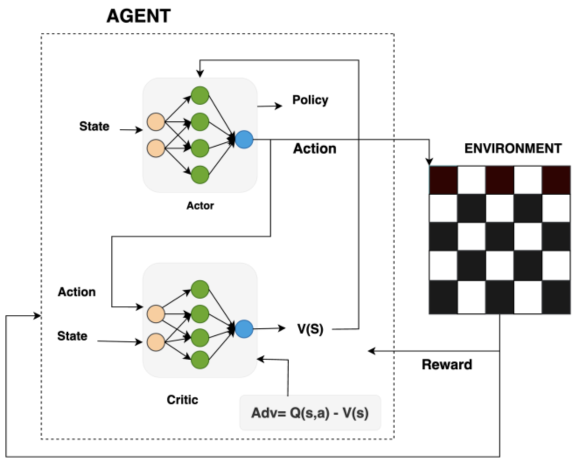

Figure 4. Actor-Critic reinforcement learning architecture

The actor selects actions based on a policy, while the critic evaluates the actions by estimating the value function. The advantage function computed as $A(s,a) = Q(s,a) - V(s)$ is used to refine the policy updates.

Despite their flexibility, Actor-Critic methods are computationally intensive, often requiring careful tuning and large amounts of training data. In educational applications like VALF, where transparency and low computational overhead are critical, the complexity of Actor-Critic models may be inappropriate. Moreover, their reliance on deep neural networks limits explainability—a key requirement when designing AI tools for learning-impaired students.

Instead, VALF opts for Cellular Learning Automata (CLA), which offers transparent, grid-based learning suited to symbolic tasks like character construction, tone placement, and phonetic association in Vietnamese language learning.

2.7 Proximal Policy Optimization (PPO)

Proximal Policy Optimization (PPO) is a modern reinforcement learning algorithm that builds on the advantages of policy gradient methods while addressing their stability and efficiency issues. Introduced by Schulman et al. [20], PPO has become one of the most widely used algorithms in reinforcement learning due to its simplicity, robustness, and state-of-the-art performance across various domains.

PPO belongs to the class of Actor-Critic methods but improves policy updates by limiting the change in policy between successive iterations, thereby avoiding performance collapse due to overly large updates. Unlike Trust Region Policy Optimization (TRPO), which relies on complex second-order optimization [21], PPO introduces a clipped surrogate objective:

$$L^{\text{CLIP}}(\theta) = E[\min(r_t(\theta)\nabla_\theta, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\nabla_\theta)] \text{(5)}$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$, and ε is a hyperparameter controlling the update step size (typically 0.1 to 0.3). The clipping mechanism ensures conservative updates, improving training stability without compromising performance. PPO has been successfully applied to a wide range of tasks, including robotics, continuous control, and video game playing. It played a central role in the OpenAI Five system that achieved expert-level performance in Dota 2 [22], and it is also the default algorithm in many reinforcement learning libraries such as Stable-Baselines3 and OpenAI Baselines [23].

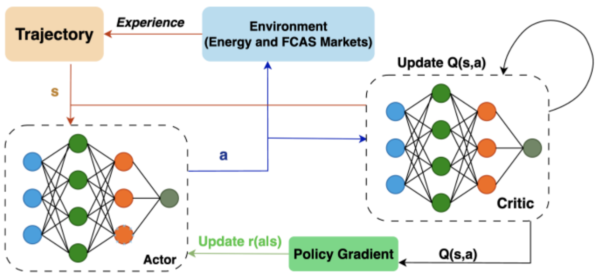

Figure 5. Workflow of Proximal Policy Optimization (PPO)

The actor network samples actions from states to interact with the environment, while the critic estimates the Q-value to guide policy gradients. The advantage function is used to stabilize policy updates.

Despite its performance, PPO still inherits some limitations of policy gradient methods, such as sample inefficiency and difficulty in interpretability. For educational applications like VALF, where model transparency and low-resource execution are priorities, PPO's reliance on neural network policies and the abstract nature of its update rules make it less ideal.

VALF, therefore, adopts Cellular Learning Automata (CLA), which enables localized, rule-based learning better suited to structured, symbol-rich educational contexts like Vietnamese orthography.

2.8 Deep Deterministic Policy Gradient (DDPG)

Deep Deterministic Policy Gradient (DDPG) is an off-policy, model-free reinforcement learning algorithm designed for environments with continuous action spaces. Introduced by Lillicrap et al. [24], DDPG extends the deterministic policy gradient (DPG) method [25] by integrating deep neural networks and incorporating key techniques from Deep Q-Networks (DQN), such as experience replay and target networks.

Unlike stochastic policy gradient methods that output action distributions, DDPG learns a deterministic policy $\mu(s;\theta)$ that directly maps states to specific actions. The algorithm employs two neural networks:

- **Actor network** $\mu(s;\theta)$: selects the action given a state.
- **Critic network** $Q(s,a;\,w)$: estimates the Q-value for a state-action pair.

The actor is updated using the deterministic policy gradient:

$$\nabla_\theta J \approx E\left[\nabla_a Q(s,a;w)|_{a=\mu(s)} \cdot \nabla_\theta \mu(s;\theta)\right] \quad (6)$$

To improve learning stability, DDPG uses target networks μ' and Q', which are soft copies of the actor and critic networks and are updated slowly using a Polyak averaging technique:

$$\theta' \leftarrow \tau\theta + (1-\tau)\theta' \; (7)$$

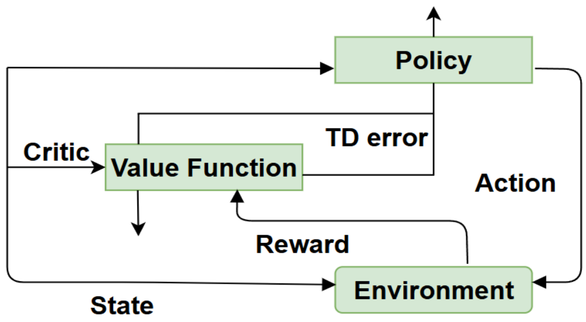

Figure 6. Architecture of Deep Deterministic Policy Gradient (DDPG)

The actor (policy) network outputs deterministic actions, while the critic (value function) evaluates their quality using temporal-difference (TD) error. Updates are stabilized with target networks.

DDPG has demonstrated strong performance in continuous control problems such as robotic manipulation, autonomous driving, and physics simulations. It has also laid the foundation for more advanced algorithms, including Twin Delayed DDPG (TD3) [26] and Soft Actor-Critic (SAC) [27].

However, DDPG is known to be sensitive to hyperparameters and exploration strategies, often requiring noise injection (e.g., Ornstein-Uhlenbeck process) for effective exploration. Its black-box neural architecture and training complexity make it less suitable for applications requiring transparency and low computational cost.

In educational settings like VALF, where interpretability, robustness, and symbolic traceability are essential, DDPG's deep and opaque learning process poses challenges. Instead, VALF adopts Cellular Learning Automata (CLA) to achieve transparent, rule-based, and symbol-sensitive reinforcement learning adapted to Vietnamese language education.

Research Gap and Motivation

Despite extensive studies in CALL for general language learning and reinforcement learning in special education, several significant gaps remain. First, no published studies have systematically investigated the application of reinforcement learning, particularly CLA-based adaptive systems, to the unique challenges of Vietnamese language learning. Second, while frameworks such as RALF have demonstrated the potential of CLA in dyslexia interventions, their extension to tonal languages remains unexplored.

This research directly addresses these gaps. The Vietnamese Adaptive Learning Framework (VALF) aims to utilize CLA's proven capabilities (as demonstrated by RALF) to tackle the phonetic, orthographic, and tonal complexities of Vietnamese, specifically targeting primary school-aged dyslexic learners.

3. THE PROPOSED MODEL: CELLULAR LEARNING AUTOMATA BASED VIETNAMESE ADAPTIVE LEARNING FRAMEWORK

This section introduces VALF model, a CLA-based reinforcement learning model specifically developed to support Vietnamese language acquisition among dyslexic primary school learners. VALF is designed as a modular system that integrates multiple sensory modalities—visual, auditory, and motor—through adaptive instruction. The framework combines Cellular Learning Automata (CLA) with a virtual pen rendering engine, synchronized audio feedback, and personalized reinforcement strategies.

3.1 Dyslexia and Vietnamese Language Learning process

Very few adaptive computer-assisted language learning (CALL) systems specifically target Vietnamese language learning for dyslexic children. Duong and Vo [28] examined the use of online platforms to foster learner autonomy and enhance reading comprehension among Vietnamese EFL learners. Their findings suggest that digital tools can significantly improve students' reading skills and promote autonomous learning behaviours. However, the platforms discussed lacked adaptive capabilities and did not employ reinforcement learning to dynamically adjust learning paths.

Recent studies on dyslexia interventions further support the use of adaptive frameworks—particularly those that offer real-time feedback and interactive components—as a means of significantly enhancing both learner autonomy and educational outcomes. For example, Smith and Hattingh [29] conducted a systematic literature review highlighting the effectiveness of assistive technologies for students with dyslexia, emphasizing the value of responsive, individualized learning systems. Similarly, Taskov and Dushanova [30] demonstrated neurological benefits of visual training interventions for learners with dyslexia, reinforcing the importance of multimodal engagement in adaptive platforms. These insights suggest that integrating adaptive reinforcement learning (RL)-based technologies specifically addressing the linguistic structure and tonal complexity of Vietnamese could significantly advance research and development in this domain. Doing so would allow for more inclusive educational technologies tailored to the cognitive profiles of learners with dyslexia in tonal language environments.

3.2 Word Training

The VALF system implements an adaptive word training mechanism based on Cellular Learning Automata (CLA) to support Vietnamese script acquisition among dyslexic primary school learners. This module simulates stroke-by-stroke character rendering using a virtual pen, synchronizing visual cues with auditory pronunciation to reinforce the phonological and orthographic structure of Vietnamese words.

Algorithm: CLA- Vietnamese Word Generation(VWG)

***Inputs:* Previous character, Main tone marks, Extension of characters as P,M,E respectively.**

***Outputs:* Graphical Shape of the Vietnamese characters**
begin

Define CA(state-set, current state,M,Prev.C,Next.C,P);

Select relevant feature in M based on the current state

Move forward based on P;

Activate the cell near to the start corordinate in M
while(an active state exists) do
choose an action according to the probabaility vector
announce the reward / penalty of the action;
If (not the end of character) then,

Go ahead by filling the corresponding feature pixel;
end while
end.

3.2.1 Character and Tone Marker Generation

Vietnamese script is built from Latin letters modified by diacritical tone marks. Each syllable consists of a consonant–vowel–tone triplet, where the vowel carries one of six tones: ngang (level, unmarked), sắc (acute), huyền (grave), hỏi (hook above), ngã (tilde), or nặng (dot below) [31]. The CLA model initializes a two-dimensional lattice in which each cell may be active (drawing), inactive (background), or designated for tone overlay. Characters are deconstructed into sequential strokes, and CLA agents iteratively learn optimal stroke sequences through reinforcement signals guided by proximity to a reference template [32].

Tone markers, being non-linear visual components, are handled separately. Once the base letter is completed, the appropriate tone is rendered as an overlay above, below, or beside the vowel depending on the tone type. This layered rendering strategy enhances clarity and reduces perceptual confusion—critical for dyslexic learners [31].

3.2.2 Virtual Pen Simulation

A virtual pen simulates the motor dynamics of handwriting by traversing the CLA grid in real-time. Each movement (up, down, left, right) is decided by localized feedback from the environment. Correct actions are rewarded, while erroneous ones are penalized, reinforcing accurate stroke formation [32]. This animated pen helps learners internalize stroke order and letter structure through visual-motor coupling. Prior research indicates that tracing letters enhances memory encoding and activates multiple neural pathways in reading acquisition [33].

3.2.3 Pronunciation Integration

As characters are rendered, VALF concurrently plays the word's pronunciation using a built-in speech engine. When the entire syllable is formed, the full audio with appropriate tone inflection is replayed. This multimodal feedback integrates auditory and visual pathways, reinforcing the association between graphemes and phonemes. For dyslexic learners, such synchrony between sound and visual input enhances recall and comprehension [34], [35].

3.2.4 Procedural Learning Loop

The training sequence in VALF follows a four-step loop:

1. Display a target image or word prompt.
2. Activate CLA-driven rendering of the character.
3. Synchronize stroke rendering with audio pronunciation.
4. Log learner response time, accuracy, and errors for adaptive feedback.

This structure promotes both passive observation and active participation. By coupling sensory channels—visual, auditory, and kinaesthetic—VALF aims to develop robust letter-sound associations and spatial awareness essential for literacy.

In summary, VALF generates each Vietnamese syllable through an incremental process: base characters are rendered stroke-by-stroke via CLA, tone markers are added in a secondary phase, and audio is synchronized throughout. This structured, multisensory approach aligns with evidence-based dyslexia intervention methods and supports enhanced learning outcomes in Vietnamese language acquisition.

3.3 Lesson Selection

The lesson structure in VALF is aligned with the Vietnamese national primary school curriculum for Grades 1 to 3, ensuring that instructional content is age-appropriate and pedagogically relevant [36]. Each unit corresponds to a thematic module—such as family, animals, or school life—drawn from the official

Vietnamese language textbooks. These units introduce a progression of linguistic complexity, starting with basic consonants and vowels, followed by tone mark application, and eventually full syllable composition and simple sentence construction. This alignment ensures systematic coverage of all literacy skills required in the early primary years [36], [37].

Each lesson integrates multiple learning modalities to support dyslexic learners:

- **Contextual Imagery and Stroke Animation:** Lessons begin with a relevant contextual image, followed by the animated rendering of Vietnamese words using a CLA-driven virtual pen. The system draws each letter and tone mark stroke-by-stroke along a horizontal writing line, helping students internalize character structure and tone placement [37].

- **Vocabulary Lists and Sentences:** The lesson interface displays vocabulary words and short example sentences. Key graphemes and tones are visually emphasized. Accompanying English translations or pictorial cues aid comprehension. Learners can tap any word or sentence to hear its correct pronunciation, supporting auditory reinforcement of written content [37].

- **Synchronized Audio Feedback:** As each character is procedurally generated, the system plays synchronized audio of the corresponding sound. This tightly couples the learner's visual and motor attention with auditory feedback, strengthening grapheme-phoneme associations [37].

VALF employs a **mastery-based progression system**. Learners must first complete the instructional lesson before accessing associated exercises or quizzes. These components remain locked until the required content has been reviewed. To advance to the next unit, students must achieve a passing score (e.g., minimum accuracy threshold), which ensures comprehension before moving forward [37]. The system records key metrics such as response accuracy, latency, and attempt frequency. Based on this data, VALF can recommend repetition of weak areas or permit faster progression for mastered skills [37].

The interface is designed for ease-of-use and accessibility:

- **Bilingual Support:** A toggle allows switching between Vietnamese and English labels, instructions, and prompts—beneficial for young learners and bilingual classrooms [37].

- **Iconic Navigation:** Interactive functions (e.g., lessons, games, quizzes) are represented using large, intuitive icons such as books, speakers, and pencils. This visual format minimizes cognitive load and enhances usability [37].

- **Dyslexic-Friendly Design:** The UI uses clean sans-serif fonts, wide spacing, and high-contrast colors. Text instructions are supplemented with audio narration. The interface avoids excessive textual content, instead relying on imagery and speech, aligning with best practices in dyslexia support [37].

In summary, VALF's lesson selection system promotes structured, multimodal, and accessible learning. It leverages curriculum alignment, progressive unlocking, and real-time learner adaptation to support reading development in dyslexic children.

3.4 VALF process as a Game Task

In VALF, each learning task is implemented as a small interactive game to enhance engagement and promote multisensory learning. Learners participate in activities such as picture-to-word matching, drag-and-drop assembly of syllables or words, and multiple-choice quizzes focused on phonics and vocabulary. These mini-games are scored, often include a countdown timer, and reward players with stars or points for correct responses, transforming routine practice into a playful and goal-oriented challenge [38].

Real-time feedback is a core component of the exercise design. Correct answers are immediately reinforced through visual highlights and celebratory audio cues, while incorrect responses prompt instant correction or encouragement (e.g., color change with a gentle prompt to retry) [38]. To support learners with dyslexia, every textual or visual element in the exercise is accompanied by synchronized audio narration. Letters and words are pronounced aloud when displayed or clicked, and game outcomes are reinforced through distinct success or failure sounds [39]. This combination of visual, auditory, and interactive feedback has been shown to significantly aid phonological decoding by strengthening the link between graphemes and phonemes [39] [40].

VALF further enhances effectiveness by adapting the difficulty level dynamically based on learner performance. Each exercise functions within a CLA-based reinforcement learning loop: if a student consistently answers correctly and rapidly, the system interprets this as a positive signal and increases the challenge—e.g., by introducing unfamiliar vocabulary, reducing response time, or adding more distractors. Conversely, if the learner struggles, VALF adapts by repeating simpler exercises, offering visual cues, or providing additional hints [41] [38]. This adaptive scaffolding approach ensures that learners remain within their zone of proximal development, avoiding both boredom from under-stimulation and frustration from excessive difficulty.

Evaluations of similar educational game systems show that learners enjoy this format and benefit from difficulty calibration that responds to their real-time performance [41]. Teachers and students report higher motivation when tasks are gamified and scaled appropriately [38] [41]. Thus, VALF's game-based exercises combine the benefits of instant correction, audio-visual reinforcement, and performance-driven personalization—providing a motivating, accessible, and effective platform for Vietnamese language learning in dyslexic students.

3.5 Multimedia Exam

VALF integrates a multimedia-based exam system that evaluates learner progress using a combination of text, image, and audio formats to ensure accessibility for dyslexic students. The exam includes four-option multiple-choice questions, picture–word association tasks, and listening-based comprehension items. For example, in a picture–word task, the learner is shown an image and asked to choose the corresponding Vietnamese word from a set of textual options—an approach proven effective for reinforcing vocabulary and lexical recognition [42].

Each question-and-answer choice is supported by synchronized audio narration. The system plays the spoken version of the question and its corresponding options while displaying the text and visuals. This dual-mode delivery leverages multisensory learning principles, allowing learners to process information through both visual and auditory channels [42].

To maintain an optimal level of challenge, the exam is adaptive. Question difficulty dynamically adjusts in response to learner performance. Correct responses lead to progressively harder questions, while incorrect responses result in easier follow-ups. This ensures that the test remains appropriate to the learner's ability while promoting gradual skill advancement [43].

Feedback is delivered in real time. Upon selecting an answer, the learner receives immediate confirmation. Correct choices are reinforced through audio repetition or visual highlights, while incorrect answers prompt the display of the correct response or a hint. This immediate corrective feedback is essential for dyslexic learners, enhancing their awareness of errors and supporting the development of accurate phonological decoding skills [42]. Spoken feedback is especially valuable, as it has been shown to significantly improve recognition, attention, and retention in students with reading difficulties [44].

All responses are scored and recorded automatically. At the conclusion of each exam session, VALF generates a summary report detailing overall accuracy, response times, and common error types. This performance data is fed back into the CLA-based learning engine to inform adaptive content delivery and track long-term learning progress.

By combining multiple question types, synchronized audio narration, real-time adaptive difficulty, and personalized feedback, VALF's multimedia exam provides an effective and inclusive assessment tool tailored to the needs of Vietnamese primary learners with dyslexia.

4. CONCLUSION AND FUTURE DIRECTIONS

In this research, the proposed model VALF as a learning framework for Vietnamese students with dyslexia to improve their reading literacy. It will provide a complete straight forward platform for learning alphabets along with tone pronunciations in Vietnamese language. VALF generates each character by a simple font, through reinforcement process by considering the Vietnamese writing method such as tone symbols, right-to-left directions, and exact alphabets usage in various forms. Finally, this flexible VALF model can also use to prepare accurate style of letters, various writing speeds along with corrosiveness and helps in improving good reading skills for primary Vietnamese children. In the future version of this article, real time implementation of the proposed VALF model in gaming platform with more interactive features for the dyslexia children were planned to execute.

5. REFERENCES

[1] Minoofam, S. A. H., Bastanfard, A., & Keyvanpour, M. R. (2022). RALF: An adaptive reinforcement learning framework for teaching dyslexic students. Multimedia Tools and Applications, 81(5), 6389–6412.

[2] Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement learning, fast and slow. Trends in Cognitive Sciences, 23(5), 408–422.

[3] Carlotto, T., & Jaques, P. A. (2016). The effects of animated pedagogical agents in an English-as-a-Foreign-Language learning environment. International Journal of Human-Computer Studies, 95, 15–26.

[4] Rello, L., & Baeza-Yates, R. (2016). The effect of font type on screen readability by people with dyslexia. ACM Transactions on Accessible Computing (TACCESS), 8(4), 1–33.

[5] Rubio, G., Navarro, E., & Montero, F. (2014). APADYT: a multimedia application for SEN learners. Multimedia Tools and Applications, 71(3), 1771–1802.

[6] Peng, X., Chen, H., Wang, L., & Wang, H. (2018). Evaluating a 3-D virtual talking head on pronunciation learning. International Journal of Human-Computer Studies, 109, 26–40.

[7] Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine Learning, 8(3–4), 229–256.

[8] Watkins, C. J. C. H. (1989). Learning from Delayed Rewards (PhD thesis).

[9] Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. Machine Learning, 8(3-4), 279–292.

[10] Riedmiller, M. (2005). Neural fitted Q iteration–first experiences with a data efficient neural reinforcement learning method. In European Conference on Machine Learning (pp. 317–328).

[11] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

[12] Lin, L.-J. (1992). Self-improving reactive agents based on reinforcement learning, planning and teaching. Machine Learning, 8(3–4), 293–321.

[13] Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double Q-learning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 30, No. 1).

[14] Wang, Z., Schaul, T., Hessel, M., van Hasselt, H., et al. (2016). Dueling network architectures for deep reinforcement learning. In Proceedings of the 33rd International Conference on Machine Learning (pp. 1995–2003).

[15] Schaul, T., Quan, J., Antonoglou, I., & Silver, D. (2015). Prioritized experience replay. arXiv preprint arXiv:1511.05952.

[16] Konda, V. R., & Tsitsiklis, J. N. (2000). Actor-Critic algorithms. In Advances in Neural Information Processing Systems (pp. 1008–1014).

[17] Mnih, V., Badia, A. P., Mirza, M., Graves, A., et al. (2016). Asynchronous methods for deep reinforcement learning. In Proceedings of the 33rd International Conference on Machine Learning (pp. 1928–1937).

[18] Lillicrap, T. P., Hunt, J. J., Pritzel, A., et al. (2015). Continuous control with deep reinforcement learning.

[19] Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In Proceedings of the 35th International Conference on Machine Learning (pp. 1861–1870).

[20] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms.

[21] Schulman, J., Levine, S., Abbeel, P., Jordan, M., &

Moritz, P. (2015). Trust region policy optimization. In International Conference on Machine Learning (pp. 1889–1897).

[22] OpenAI (2019). OpenAI Five. https://openai.com/blog/openai-five/.

[23] Raffin, A., Hill, A., Gleave, A., Kanervisto, A., & Ernestus, M. (2021). Stable-Baselines3: Reliable reinforcement learning implementations. Journal of Machine Learning Research, 22(268), 1–8.

[24] Lillicrap, T. P., Hunt, J. J., Pritzel, A., et al. (2015). Continuous control with deep reinforcement learning.

[25] Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., & Riedmiller, M. (2014). Deterministic policy gradient algorithms. In Proceedings of the 31st International Conference on Machine Learning (pp. 387–395).

[26] Fujimoto, S., Van Hoof, H., & Meger, D. (2018). Addressing function approximation error in actor-critic methods. In International Conference on Machine Learning (pp. 1587–1596).

[27] Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In Proceedings of the 35th International Conference on Machine Learning (pp. 1861–1870).

[28] Duong, T. H. D., & Vo, T. T. H. (2024). Fostering Vietnamese EFL learners' learner autonomy and reading comprehension ability through online platforms. European Journal of English Language Teaching, 9(4), 113–131.

[29] Smith, C., & Hattingh, M. J. (2020). Assistive technologies for students with dyslexia: A systematic literature review. In International Conference on Innovative Technologies and Learning, Springer, pp. 504–513.

[30] Taskov, T., & Dushanova, J. (2021). Small-World Propensity in Developmental Dyslexia After Visual Training Intervention. In Intelligent Computing. Springer, pp. 233–258.

[31] D. Lam, *Vietnamese Typography*. [Online]. Available: https://vietnamesetypography.com/.

[32] S. A. H. Minoofam, A. Bastanfard, and M. R. Keyvanpour, "RALF: An adaptive reinforcement learning framework for teaching dyslexic students," *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 6389–6412, 2022. doi: 10.1007/s11042-021-11806-y.

[33] International Dyslexia Association, "Understanding dyslexia and the reading brain in kids," 2020. [Online]. Available: https://dyslexiaida.org.

[34] LDAU, "Multisensory learning and its effects on dyslexia," 2021. [Online]. Available: https://ldau.org

[35] G. Rubio, E. Navarro, and F. Montero, "APADYT: A multimedia application for SEN learners," *Multimedia Tools and Applications*, vol. 71, no. 3, pp. 1771–1802, 2014. doi: 10.1007/s11042-012-1304-9.

[36] Nord Anglia Education, "Vietnamese National Curriculum," [Online]. Available: https://www.nordangliaeducation.com/.

[37] S. A. H. Minoofam, A. Bastanfard, and M. R. Keyvanpour, "RALF: An adaptive reinforcement learning framework for teaching dyslexic students," *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 6389–6412, 2022. doi: 10.1007/s11042-021-11806-y.

[38] A. Ahmad and M. S. Ahmed, "An Interactive Learning Platform for Dyslexic Children using Multimedia and Game-based Techniques," *International Journal of Interactive Mobile Technologies*, vol. 14, no. 6, pp. 127–139, 2020. [Online]. Available: https://www.warse.org.

[39] CAST, "Audio-Supported Reading for Students with Learning Disabilities," *CAST Research Reports*, 2021. [Online]. Available: https://aem.cast.org.

[40] A. Smith and C. Hattingh, "Assistive technologies for students with learning disabilities: A review," *Disability and Rehabilitation: Assistive Technology*, vol. 15, no. 1, pp. 11–24, 2020. doi: 10.3109/17483107.2019.1573840.

[41] H. Yildirim and E. Surer, "Design and Evaluation of Adaptive Serious Games for Dyslexic Children," *PubMed Central*, 2021. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov.

[42] A. K. Singh and R. D. Rishi, "An Adaptive E-Learning Framework for Children with Dyslexia," *PubMed Central*, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc.

[43] A. T. Albano, "The Future of Adaptive Testing," *Credentialing Insights*, vol. 31, no. 2, pp. 10–15, 2023. [Online]. Available: https://www.credentialinginsights.org.

[44] Acapela Group, "Text-to-speech solutions for dyslexia," *Acapela Group*, 2022. [Online]. Available: https://www.acapela-group.com.