

Article

Neural Signatures of Speed and Regular Reading: A Machine Learning and Explainable AI (XAI) Study of Sinhalese and Japanese

Thishuli Walpola ¹ , Namal Rathnayake ^{2,*} , Hoang Ngoc Thanh ^{1,3} , Niluka Dilhani ¹  and Atsushi Senoo ^{1,*} 

¹ Department of Radiological Sciences, School of Human Health Sciences, Tokyo Metropolitan University, Tokyo 116-0012, Japan; 24997707@ed.tmu.ac.jp (T.W.); 23997706@ed.tmu.ac.jp (H.N.T.); 25997706@ed.tmu.ac.jp (N.D.)

² Advanced Institute for Marine Ecosystem Change (WPI-AIMEC), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama 236-0001, Japan

³ Department of Radiology, Hue University of Medicine and Pharmacy, Hue 53000, Vietnam

* Correspondence: namalr@jamstec.go.jp (N.R.); senoo@tmu.ac.jp (A.S.)

Abstract

Reading speed is hypothesized to have distinct neural signatures across orthographically diverse languages, yet cross-linguistic evidence remains limited. We investigated this by classifying speed readers versus regular readers among Sinhalese and Japanese adults ($n = 142$) using task-based fMRI and 35 supervised machine learning classifiers. Functional activation was extracted from 12 reading-related cortical regions. We introduced Fuzzy C-Means (FCM) clustering for data augmentation and Shapley additive explanations (SHAP) for model interpretability, enabling evaluation of region-wise contributions to reading speed classification. The best model, an FT-TABPFN network with FCM augmentation, achieved 81.1% test accuracy in the Combined cohort. In the Japanese-only cohort, Quadratic SVM and Subspace KNN each reached 85.7% accuracy. SHAP analysis revealed that the angular gyrus (AG) and inferior frontal gyrus (triangularis) were the strongest contributors across cohorts. Additionally, the anterior supra marginal gyrus (ASMG) appeared as a higher contributor in the Japanese-only cohort, while the posterior superior temporal gyrus (PSTG) contributed strongly to both cohorts separately. However, the posterior middle temporal gyrus (PMTG) showed less or no contribution to the model classification in each cohort. These findings demonstrate the effectiveness of interpretable machine learning for decoding reading speed, highlighting both universal neural predictors and language-specific differences. Our study provides a novel, generalizable framework for cross-linguistic neuroimaging analysis of reading proficiency.

Keywords: reading speed classification; fMRI; explainable AI (XAI); SHAP analysis; cross-linguistic neuroimaging



Academic Editor: Yu-Cheng Wang

Received: 6 December 2025

Revised: 13 January 2026

Accepted: 15 January 2026

Published: 21 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Functional Magnetic Resonance Imaging (fMRI) has emerged as a non-invasive technique to read out signals from the brain and interpret them to understand functional tasks [1]. The signals extracted from blood-oxygen-level-dependent (BOLD) fMRI acquisition have been a source of physiological information analysis in the vast field of neuroimaging [2]. Within this scope, the task of reading, and especially speed reading, has become a significant concern for researchers recently [3]. However, the quantitative

extraction of signal values from the BOLD signal and its implications for focusing on the speed-reading function are yet to be fully explored; hence, research regarding using eye-tracking methodologies to examine how readers allocate visual attention across text [4,5], the semantic processing of the brain, and the comprehension-based rating for reading [6,7] have been conducted primarily for qualitative interpretations in previous research studies.

Reading is a complex cognitive process engaging distributed neural networks responsible for visual processing, phonological decoding, and semantic integration [8]. Core regions include the visual word-form area of the ventral occipitotemporal cortex (VWFA), inferior frontal gyrus (IFG), superior temporal gyrus (STG), angular gyrus (AG), and middle temporal gyrus (MTG), which facilitate the translation of visual symbols into linguistic representations necessary for comprehension [9,10]. Literacy acquisition fundamentally reorganizes these networks, enhancing early visual processing in the occipital cortex, increasing specialization within the VWFA for orthographic input, and strengthening connectivity between phonological and graphemic representations [11,12]. Reading proficiency is supported by the integration of visual input with phonological and semantic systems, and the efficient functioning of these networks is a hallmark of skilled readers [13]. Longitudinal studies reveal that as reading proficiency develops, there is increased left-lateralization of activation within language-related regions, reflecting specialization for rapid decoding and semantic processing [14]. Executive function regions, including the prefrontal cortex, contribute to maintaining attention and working memory during reading, further supporting reading fluency [15].

Fluent reading depends on coordinated activity across ventral and dorsal brain pathways. The ventral visual–lexical stream, centered in the left ventral occipitotemporal cortex, supports rapid recognition of written forms and efficient access to lexical representations. With increasing reading experience, this pathway becomes more specialized and automatic [12,16]. In contrast, dorsal temporo-parietal and inferior frontal regions support phonological decoding, articulatory planning, and executive control during reading. The degree to which these regions are engaged varies with task demands and reader proficiency [17,18]. Neuroimaging studies of reading development and training show that higher reading fluency is associated with changes in how these systems interact. Faster reading typically involves reduced reliance on serial phonological processing and more efficient ventral-stream processing [19–21]. Studies using sentence-level and naturalistic reading tasks further demonstrate that fluent reading relies on dynamic interactions over time. These interactions link visual word-form regions with higher-order semantic and language networks during continuous comprehension [22–24].

1.1. Speed Reading and Neuroplasticity

Recent neuroimaging studies have delineated the fundamental reading network and demonstrated that reading speed modulates activity in the occipitotemporal, temporo-parietal, and inferior frontal regions. Traditional stimulus-rate and reading acceleration paradigms reveal systematic rate effects in word reading involving temporal and frontal language regions [20,25,26]. Furthermore, speed-reading interventions have been associated with reduced reliance on slower phonological pathways and altered engagement of language circuits [21]. Recent studies into naturalistic or connected-text reading underscore that fluency entails synchronized interactions between ventral visual–lexical areas and advanced language or semantic systems [24,27].

Speed reading, whether achieved naturally or through training, appears to modulate activity within these core reading networks. fMRI studies demonstrate that as reading speed increases, activation within language-related regions, including the left STG, IFG, and anterior temporal regions, often decreases, suggesting reduced reliance on phono-

logical and syntactic processing during rapid reading [21,28]. Conversely, visuospatial regions such as the right intraparietal sulcus exhibit increased activation, reflecting the enhanced visuospatial processing demands associated with faster eye movements and broader perceptual spans during speed reading [28]. Speed-reading training can induce functional reorganization within language and attention networks. For instance, a study demonstrated that after speed-reading training, participants exhibited decreased resting-state connectivity between Broca's area and its right-hemisphere homolog, potentially reflecting reduced subvocalization during reading, a commonly cited mechanism for increased reading speed [29]. Additionally, training may enhance connectivity with auditory regions while decreasing coupling with memory-associated areas, supporting more direct semantic access during reading [29]. In studies of children with reading difficulties, reading acceleration training has been shown to enhance left-hemisphere activation, indicating normalization of reading-related neural circuits [14,15]. Notably, training can also engage compensatory mechanisms involving frontal executive networks, suggesting that effective speed-reading training leverages both typical and compensatory pathways to support reading fluency [15,30]. Recent research focusing on native Japanese readers undergoing speed-reading tasks has demonstrated complex connectivity within the occipitotemporal network, with increased effective connections from the inferior occipital gyrus to the VWFA and modulation effects linked to reading speed [31]. These findings underscore the dynamic reorganization of reading networks during speed reading, emphasizing the roles of both visual processing pathways and higher-order language regions in facilitating increased reading speeds. An fMRI study examining the differential effects of word presentation rates (reading acceleration) on the brain activity patterns of dyslexic readers found no significant differences in activation patterns between the two groups during 'slow' reading of sentences and single words; meanwhile, 'fast' reading was associated with increased activation in visual areas among normal readers [20].

The previous study results indicate that once reading has become fully automated, fast and fluent reading is primarily controlled by the ventral reading system [32]. The left hemisphere (LH) ventral system includes the clusters in the left inferior frontal gyrus pars triangularis (IFGT) and orbitalis and the posterior middle temporal gyrus (PMTG), as well as the left anterior temporal lobe and angular gyrus (AG) [30,33]. Furthermore, the left IFGT supports semantic control and executive regulation of retrieval. Speed reading typically shifts the balance of control demands by reducing phonological assembly and increasing rapid integration, creating systematic group differences that a classifier can exploit [34]. AG, being a multimodal semantic integration hub, is repeatedly implicated in reading and comprehension, which means that proficient reading across languages tends to recruit this territory. Differences in semantic integration strategies, which show a preference for speed in speed readers and a focus on deep comprehension in regular readers, plausibly separate speed readers from regular readers [35,36]. In typical adult readers, increasing the word presentation rate (150 → 240 → 600 wpm) produced greater activation in the left MTG (peak −50, −46, 8), alongside visual cortices [27,37,38]. However, recent studies of connectivity patterns reveal that this semantic integration hub tends to show stronger functional coupling with orthographic areas (e.g., VWFA) in slower readers, consistent with deeper semantic reliance, rather than in speed readers [24].

1.2. Cross-Linguistic Context

Cross-linguistic fMRI studies also show that the universal reading network is modulated by script properties. In Japanese, for example, comparisons of Kana and Kanji (and L2 script learning) reveal script-dependent shifts in occipito-temporal and temporo-parietal engagement, consistent with differences in orthography–phonology

mapping demands [39–41]. However, these studies typically use univariate contrasts or group-level statistics. These studies rarely test whether multivariate neural features can generalize across orthographies to identify reader types, nor do they quantify which regions drive such classification in an interpretable way.

Orthographic and phonological differences across writing systems influence neural activation during reading, with cross-linguistic fMRI studies revealing both shared and language-specific processing routes for alphabetic versus logographic scripts [42]. The script relativity hypothesis further posits that the visual characteristics of a writing system influence cognitive strategies and brain networks, a view supported by focused research agendas on script effects [43]. Bilingual neuroimaging experiments demonstrate the dynamic recruitment of cognitive control regions during language membership control tasks, indicating that ML models of neural data must accommodate cross-language interactions [44]. Despite this progress, a paucity of ML-driven studies remains on under-represented scripts, such as Sinhala, highlighting a critical gap for future work.

Japanese uses a mixed script combining syllabaries, where each character represents a full syllable, alongside logographic Kanji, borrowed from Chinese [45]. In contrast, Sinhala employs an abugida, also called an alphasyllabary, rooted in the Brahmi script [46]. This leads to distinctly different orthographic architectures. Japanese treats each syllabic unit as an indivisible glyph, whereas Sinhala composes syllables from base consonant forms modified by appended vowel signs.

1.3. Methodological Shift: ML and XAI

Machine learning has become central to cognitive neuroscience by enabling the discovery of multivariate patterns in high-dimensional neural datasets that traditional univariate analyses often overlook [47]. Early landmark work applied ML to decode perceptual and semantic states from distributed fMRI activations using multi-voxel pattern analysis (MVPA), achieving reliable classification of object categories and mental imagery [48]. Beyond classification, supervised ML serves four primary roles in systems neuroscience: solving engineering problems (e.g., brain–computer interfaces), identifying predictive biomarkers, benchmarking simple brain models, and providing abstract computational models that inform theories of neural coding [47]. Recent advances extend ML to more complex encoding and decoding frameworks. For instance, transfer learning from NLP-task-specific representations has improved the prediction of fMRI responses to natural language stimuli [49], while deep recurrent architectures, such as LSTM-based frameworks, maintain interpretability via layer-wise relevance propagation [50].

Machine learning (ML) techniques have experienced rapid adoption in cognitive neuroscience due to their ability to detect complex, multivariate patterns in neural data that traditional univariate analyses may overlook [51,52]. However, different ML algorithms make distinct assumptions about data distributions and model complexity, resulting in their performance varying substantially across datasets [53,54]. Therefore, training and comparing multiple models, ranging from support vector machines and random forests to deep neural networks and graph-based approaches, provides robustness against algorithm-specific biases and improves the generalizability of results to new subjects and recording conditions [54]. Ensemble and multi-model frameworks further leverage the complementary strengths of individual learners, yielding more stable and accurate predictions in neuroimaging classification and regression tasks [53–55].

Alongside these theoretical advances, multivariate decoding has become an important approach for studying language and reading in the brain. Early work in multivariate pattern analysis demonstrated that information about stimulus categories is distributed across patterns of fMRI activity rather than being confined to single brain regions [56].

This insight encouraged the development of predictive models that integrate signals from multiple regions simultaneously. Subsequent semantic decoding studies showed that multivariate models can successfully predict conceptual and semantic content from brain activity, providing a direct computational link between neural responses and linguistic meaning [57]. In reading research, multivariate approaches have been applied to naturalistic story comprehension tasks. These studies demonstrate that different brain systems contribute dynamically over time during connected-text reading, allowing distinct reading subprocesses to be examined within continuous comprehension [22].

The black-box nature of many ML models has spurred the development of Explainable AI (XAI) methods that render predictions transparent, with systematic reviews cataloging post hoc techniques such as LIME and SHAP for neuroimaging applications [58]. These reviews emphasize that XAI methods vary in reliability and fidelity, underscoring the need for critical evaluation when interpreting findings in cognitive neuroscience [58]. SHAP, in particular, provides a game-theoretic approach to feature attribution, assigning each neural feature a Shapley value that quantifies its contribution to model outputs [59]. Applications of SHAP frameworks have mapped multivariate relationships between regional neuroimaging biomarkers (e.g., amyloid-PET and MRI volumes) and cognitive scores in MCI/Alzheimer's disease, revealing key modality-specific contributors to global cognitive decline [59].

Despite their predictive power, many ML models remain opaque, hindering the extraction of neurobiologically meaningful insights and limiting their translational impact [52,58]. Explainable AI (XAI) methods, such as Local Interpretable Model-agnostic Explanations (LIME) and Shapley additive explanations (SHAP), have emerged as essential tools for elucidating model decision-making processes in neuroimaging studies [59]. SHAP, in particular, offers a unified, game-theoretic framework that quantifies the contribution of each feature, such as voxel intensity or region of interest activation, to model outputs, facilitating a direct mapping of the neural substrates underlying cognitive phenomena [59]. Applications of SHAP in cognitive neuroscience span from mapping multivariate relationships between biomarkers and cognition in Alzheimer's disease [48,59] to decoding region-wise contributions in EEG-based cognitive assessment [53,60], demonstrating its utility for uncovering neurobiological mechanisms and guiding hypothesis-driven inquiries.

1.4. Research Gaps and Problem Statement

A previous fMRI study on the identification of cross-linguistic neural strategies for reading Japanese and Chinese was conducted to assess brain activation patterns in each language among native speakers [7]. Further, a meta-analysis was performed to measure the effects of age of appropriation (AoA) and second language proficiency on language representation in the bilingual brain [61]. However, these studies showcase the research gaps due to the lack of cross-language neural studies based on machine learning (ML). Although existing ML studies, such as that identifying brain regions in Alzheimer's disease [62], contain Shapley additive explanations (SHAP)-based analysis, research often lacks transparent interpretation using SHAP to understand region-specific contributions to reading speed. The changes in brain activation patterns and the connectivity pathways associated with reading speed have been studied before [21,28,31].

However, distinguishing the magnitude of changes in these patterns in different brain regions between trained speed readers and regular readers remains underexplored. Therefore, this further demonstrates that Explainable AI (XAI) methods, particularly SHAPley Additive exPlanations (SHAP), have become increasingly common in neuroimaging for interpreting predictive models [63]. SHAP has been applied to various neuroimaging problems, such as brain-age modeling and the prediction of neurodegenerative diseases,

and is often emphasized in reviews of XAI in neuroimaging [53,59]. Therefore, novelty in the present work does not lie in using SHAP itself. Rather, its theoretical contribution lies in applying interpretable ML methods to a reading-neuroscience question that conventional group-contrast fMRI cannot directly address. Specifically, this study examines whether neural features that distinguish highly fluent speed readers from regular readers generalize across orthographically distinct languages or remain script-specific. In the meantime, we use SHAP to provide transparent, region-wise contributions to model decisions. Furthermore, this study identifies which components of the reading network provide this discriminative information. By combining interpretable classification with cross-linguistic analysis, the present study moves beyond demonstrating model accuracy. It provides a principled approach for evaluating the portability and specificity of neural markers of reading fluency.

However, research using a machine learning approach to detect the differential neural contributions of brain regions in speed and regular reading has never been conducted. Therefore, it is also important to identify the underlying reasoning for each brain region that exhibits high or low activity during speed reading, which can be utilized to understand the brain regions and their contributions to the model classification.

Therefore, based on the above studies, three main research gaps have been identified. The first is that interpretable, multivariate reader-type classification in task-based reading in fMRI is rare. Most reading-speed fMRI studies quantify rate effects with univariate contrasts rather than explainable multivariate classification [20,21,25,64]. Secondly, the cross-orthography generalization is seldom tested. Cross-script studies document differences (e.g., Kana vs. Kanji in Japanese) but do not test whether a common neural signature can classify reader types across languages [39–41]. Lastly, reader heterogeneity is often treated as noise rather than as a structural factor. Differences in strategies may give rise to subtypes that affect classification and interpretability. However, soft clustering (e.g., Fuzzy C-Means) provides a data-driven approach to capturing the graded subgroup structure rather than assuming a single homogeneous population.

1.5. The Present Study: Objectives and Hypotheses

Accordingly, the present study is structured around three interrelated research questions. First, it examines whether explainable machine learning models trained on task-based fMRI features can reliably distinguish speed readers from regular readers within each language group, namely Sinhala and Japanese. Second, it evaluates whether classification performance remains above chance when the two language cohorts are combined, thereby assessing the extent to which discriminative neural features generalize across orthographically distinct writing systems. Finally, this study investigates which brain regions contribute most consistently to model predictions using SHAP-based explanations.

Therefore, the primary objective of this study is to identify individuals as speed or regular readers within each language group and across both groups combined, using a SHAP-based ML model. Specifically, the study aims to determine the direction (positive/negative) and magnitude of region-wise contributions to classification accuracy, thereby identifying the most discriminative neural features distinguishing speed and regular readers. Furthermore, the study focuses on using FCM-based features to uncover latent inter-regional relationships in reading networks and boost classification accuracy.

Therefore, based on the prior literature, this study hypothesized the following: (1) SHAP-based machine learning models will achieve accurate classification of speed readers versus regular readers, demonstrating the viability of using interpretable ML for decoding the neural correlates of reading speed. (2) The highest contribution to the model classification will be provided from the left IFGT and left AG belonging to the ventral

reading system. (3) The left PMTG region contributes less discriminative value for model classification and differentiating reader types.

2. Methods

2.1. Data Preparation

2.1.1. Participants and Study Design

This study employed a cross-sectional between-groups comparative design. A total of 142 healthy adult participants took part in this study. The cohort comprised 56 native Sinhalese and 35 native Japanese who read at regular speeds (regular readers), as well as 28 native Sinhalese and 23 native Japanese who read at a high speed (speed readers) compared to regular readers. The average reading speeds for each of the groups, along with their demographic details, are shown in Table 1 below. Each participant underwent an fMRI scanning session. The age, handedness, and sex were kept similar to minimize the inter-subject variability.

Table 1. The demographic details of the sample.

	Sinhala Readers		Japanese Readers	
	Regular	Speed	Regular	Speed
Age (Mean $\pm$ Std)	31 $\pm$ 3	30 $\pm$ 4	23 $\pm$ 6	24 $\pm$ 7
Gender (Male/Female)	28:28	14:14	26:9	16:7
Handedness (Right/Left)	56 (Right)	28 (Right)	33:2	21:2
Average Reading Speed	In words per minute (wpm)		In characters per minute (cpm)	
	240 $\pm$ 63	685 $\pm$ 175	940 $\pm$ 612	3816 $\pm$ 1537

All participants had normal or corrected-to-normal vision. Handedness was assessed in the mean laterality quotient using the Edinburgh Handedness Inventory [65]. Informed consent was obtained from all participants. This study was approved by the Human Ethics Review Committee of Human Health Sciences, Tokyo Metropolitan University (Approval No. 22022). In addition, all the methods were performed in accordance with the relevant guidelines and regulations of the Human Ethics Review Committee of Human Health Sciences, Tokyo Metropolitan University.

2.1.2. MRI Acquisition

MRI data were collected on a 3.0 T GE Signa Premier scanner (GE Healthcare, Chicago, IL, USA) using a 48-channel head coil. Functional images were acquired with a gradient-echo EPI sequence (TR = 1 s, TE = 30 ms, FA = 60°, 3 mm slices; matrix 64 $\times$ 64; FOV 220 $\times$ 220 mm). High-resolution 3D T1-weighted anatomical images were collected using an MP-RAGE sequence (TR = 2.21 s, TE = 3.34 ms, FA = 8°, 2 mm slices).

2.1.3. fMRI Task

Participants covertly read passages in their native language (Sinhala or Japanese). Sinhala passages were drawn from the Sinhala translation of *Oliver Twist*, and Japanese passages from the Japanese translation of *Gulliver's Travels*. Figure 1 illustrates the session structure: six runs, each comprising alternating 30 s rest and 60 s reading blocks (white text on a black background; visual angle: 0.46°). Texts were presented via MR-compatible goggles, and luminance was standardized (characters: 0.75 cd/m²; background: 160 cd/m²). The protocol was adapted from the study by Saito [66]. Participants confirmed that they

had either not read the assigned novel or had only read it several years earlier. Reading speed was calculated in words per minute (WPM).

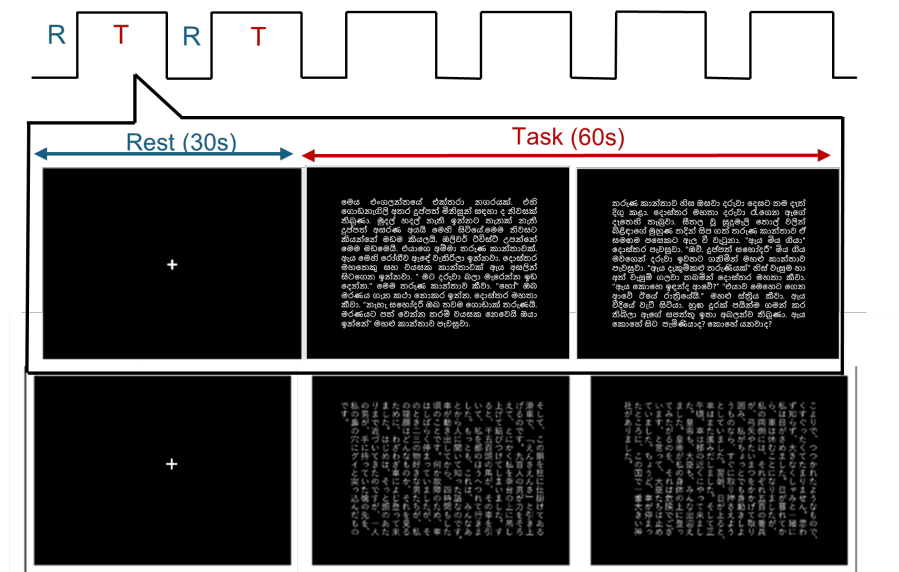


Figure 1. Block design of the covert reading fMRI task. Each session comprised six runs with alternating 30 s rest (R; fixation) and 60 s reading (T) blocks. Passages were shown in the participant's native language (Sinhala or Japanese) as white text on a black background via MR-compatible goggles.

2.1.4. fMRI Preprocessing and Analysis

All fMRI preprocessing and analysis tasks were conducted using the FMRIB Software Library (FSL, v6.0.7) in line with established protocols for task-based fMRI [67]. Standard preprocessing steps included motion correction, brain extraction, intensity normalization, temporal block segmentation, spatial normalization, and region-wise signal extraction for subsequent machine learning analysis.

Motion correction was performed using the Linear Image Registration Tool (MCFLIRT) to realign all functional volumes. A mean functional image was calculated, and the brain extraction was performed using the Brain Extraction Tool (BET) with a fractional intensity threshold of 0.3 to generate a brain mask. This mask was applied to the motion-corrected functional data to remove non-brain signals. FSLeves visual quality control (QC) was utilized to identify the outliers.

For temporal filtering and task segmentation, the fMRI time series was divided into task and rest blocks based on the experimental design using fsfMRI. Task blocks were merged into a single 4D image and averaged to create a mean task activation image. Similarly, rest blocks were processed to generate a mean rest image. Smoothing was performed with a full-width half-maximum value of 5 mm. Spatial normalization to MNI152 standard space (2 mm) was performed using FLIRT (flirt) to align the mean task and rest images to the standard template. Alignment quality was verified using FSLeves with overlaid anatomical and region of interest (ROI) masks. Furthermore, subjects were asked to remain still during the scanning time as a precaution to minimize motion artifacts.

For region-wise signal extraction, binarized ROI masks were applied using fsfMRI to compute the mean task BOLD signal within each region. The selected brain ROI masks were created using the threshold values in the default Harvard-Oxford Cortical and Subcortical Atlases in FSL (HarvardOxford-cort-maxprob-thr25-2mm.nii and HarvardOxford-sub-maxprob-thr25-2mm.nii). The signal extraction was focused on a priori regions implicated in reading fluency and accelerated speed based on previous neural studies. L/IFG, encompassing Broca's area, has been repeatedly shown to decrease activation in trained Japanese speed readers during rapid reading, reflecting reduced subvocal articulation [28].

The crucial brain regions representing the dorsal and ventral pathways of reading, which include the precentral gyrus (Prec); inferior frontal gyrus pars opercularis (IFGO); STG, responsible for phonological processing; left IFGT; left posterior middle temporal gyrus (PMTG); and AG, responsible for semantic processing, were also considered [30]. At the same time, the left anterior and posterior superior temporal gyrus (ASTG and PSTG), key phonological and semantic processing areas, exhibited diminished BOLD responses as reading speed increased [28]. The left AG has also been found to reduce activation in fast readers, correlating negatively with comprehension speed [66]. Non-language regions such as the hippocampus (Hippo) were also included based on evidence that rapid readers show attenuated engagement in working memory and memory retrieval networks [66]. The above findings, alongside comprehensive network analyses by researchers, further support the involvement of occipitotemporal connectivity in rapid reading [31]. These include dynamic interactions between posterior temporal regions and visual processing hubs, which also encompass the VWFA. Moreover, it was found that training effects in reading-impaired children have recently been linked to changes in fMRI activation, including posterior occipitotemporal regions in the vicinity of the VWFA [68].

These mean task BOLD signals were extracted from the above-selected brain regions for each participant and used as input features for machine learning classification models and SHAP-based interpretation to identify region-wise contributions to speed reading across participants.

2.2. Data Preprocessing for ML

Three distinct datasets were assembled for analysis: one collected from Sri Lanka, one from Japan, and a combined dataset encompassing both regions. Each sample was labeled as a regular reader (class 0) or a speed reader (class 1). Prior to feature extraction, we addressed class imbalance using SMOTE (see Equation (1)), generating synthetic minority-class examples to equalize the two reader categories. Subsequently, Fuzzy C-Means clustering was applied to derive soft membership features (Equation (2)), and these fuzzy-membership values were concatenated with the original predictors.

2.2.1. Sequence of Preprocessing and Leakage Prevention

To ensure the integrity of the performance estimates, all data augmentation and feature extraction steps were performed strictly within the cross-validation framework. Following the initial training–test split (or fold partition), the Synthetic Minority Over-sampling Technique (SMOTE) was applied exclusively to the training partition to equalize class distributions [69].

Similarly, the FCM clustering algorithm was fitted solely on the training data to determine optimal cluster centers. These centers were subsequently used to transform both the training and test sets into soft-membership and distance-based feature spaces. This strict nesting ensured that no information from the test set, either through class distribution or topological cluster structure, was leaked into the training phase, thereby maintaining the validity of the reported generalization metrics.

2.2.2. SMOTE for Class Balancing

To mitigate class imbalance, we employed the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic minority-class samples. For each minority sample x_i , SMOTE selects one of its k -nearest neighbors x_{nn} and creates a new sample

$$x_{\text{new}} = x_i + \lambda (x_{nn} - x_i) \quad (1)$$

where $\lambda \sim \mathcal{U}(0, 1)$ (see Equation (1)). By interpolating between existing minority points, SMOTE balances the class distribution without simply duplicating observations. The augmented dataset was then used for all subsequent model training.

2.2.3. Fuzzy C-Means Clustering Features

We applied Fuzzy C-Means (FCM) clustering to the input data to extract soft cluster-membership features. FCM minimizes the objective

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2 \quad (2)$$

where x_i is the i th sample, c_j the j th cluster center, $u_{ij} \in [0, 1]$ the membership degree of x_i in cluster j , and $m > 1$ the fuzzifier exponent (see Equation (2)). Memberships u_{ij} are updated iteratively alongside centers c_j until convergence. After fitting, the resulting membership matrix $U = [u_{ij}]$ provides C additional features per sample—each u_{ij} captures the degree to which x_i belongs to cluster j . These fuzzy-membership features were concatenated with the original predictors to enrich the feature space.

2.3. Machine Learning Models

We evaluated a diverse set of 35 supervised classifiers, organized into seven families. Hyperparameters for each model were tuned via grid search with 5-fold cross-validation, using classification accuracy as the fitness function.

2.3.1. Decision-Tree Classifiers

Decision-tree classifiers (Fine, Medium, Coarse) recursively partition the feature space by selecting the split that yields the greatest reduction in node impurity. In our implementation, impurity is measured by the Gini index

$$G = \sum_{k=1}^K p_k(1 - p_k) \quad (3)$$

where p_k is the proportion of samples of class k in a node. Each split is chosen to minimize the weighted sum of child-node Gini indices (see Equation (3)).

2.3.2. Discriminant Analysis

Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) assume Gaussian class-conditional distributions and seek projection vectors that maximize class separability. For LDA, this is achieved by maximizing the Fisher criterion

$$J(w) = \frac{w^\top S_B w}{w^\top S_W w} \quad (4)$$

where S_B and S_W are the between-class and within-class scatter matrices, respectively. QDA relaxes the shared-covariance assumption, allowing each class to have its own covariance matrix (Equation (4)).

2.3.3. Logistic Regression

Binary GLM Logistic Regression and its Efficient variant model the posterior probability via the sigmoid link $\sigma(z) = 1/(1 + e^{-z})$ and optimize the regularized negative log-likelihood (cross-entropy)

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \ln \sigma(w^\top x_i) + (1 - y_i) \ln(1 - \sigma(w^\top x_i))] \quad (5)$$

where $y_i \in \{0, 1\}$, x_i is the feature vector, and w is the weight vector (see Equation (5)).

2.3.4. Support Vector Machines (SVMs)

Linear, Polynomial (Quadratic, Cubic), and Gaussian (Fine, Medium, Coarse) SVMs solve

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \max(0, 1 - y_i(w^\top \phi(x_i) + b)) \quad (6)$$

where $\phi(\cdot)$ is the feature map induced by the chosen kernel and C the regularization parameter. This hinge-loss formulation maximizes the margin between classes (Equation (6)).

2.3.5. Naïve Bayes

Gaussian and Kernel Naïve Bayes classifiers assume feature independence and compute posterior probabilities via

$$p(C_k | x) = \frac{p(C_k) \prod_j p(x_j | C_k)}{\sum_\ell p(C_\ell) \prod_j p(x_j | C_\ell)} \quad (7)$$

where Gaussian Naïve Bayes estimates each $p(x_j | C_k)$ as a normal density and Kernel Naïve Bayes uses a nonparametric kernel density estimate (Equation (7)).

2.3.6. Instance-Based Learners (k-NN)

k-nearest neighbors classifiers (Fine, Medium, Coarse, Cosine, Cubic, Weighted, Subspace) assign class labels by majority vote among the k closest training points under a distance metric d . For Euclidean-based variants:

$$d(x, x_i) = \|x - x_i\|_2 \quad (8)$$

and for weighted k-NNs, votes are inversely weighted by $d(x, x_i)$. Subspace k-NNs operate on random feature subsets to improve robustness (see Equation (8)).

2.3.7. Ensemble Methods

Boosted Trees and RUSBoosted Trees iteratively fit weak learners to re-weighted data by minimizing the exponential loss

$$L = \frac{1}{N} \sum_{i=1}^N e^{-y_i F_{t-1}(x_i)} h_t(x_i) \quad (9)$$

where F_{t-1} is the ensemble prediction after $t - 1$ rounds and h_t the new tree's output. Bagged Trees reduce variance by averaging predictions of multiple trees trained on bootstrap samples. Subspace Discriminant ensembles apply Discriminant Analysis within random subspaces (Equation (9)).

2.3.8. Artificial Neural Networks

Narrow, Medium, Wide (single hidden layer), Bilayered, and Trilayered feed-forward networks are trained via backpropagation to minimize the cross-entropy loss

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K y_{i,k} \ln \hat{y}_{i,k} \quad (10)$$

where $\hat{y}_{i,k}$ is the softmax output for class k on sample i , and network depth and width control model capacity (see Equation (10)).

2.3.9. Kernelized Models and FT-TABPFN

Kernelized SVM and Logistic Regression apply the “kernel trick” within their respective hinge-loss or log-likelihood frameworks, replacing inner products with $K(x_i, x_j)$. The FT-TABPFN (Fast Tabular Prior-Data Fitted Network) approximates few-shot Bayesian classification by learning a transformer network that predicts the posterior predictive distribution enabling rapid adaptation to new tabular tasks (Equation (11)).

$$p(y | x, D) = \int p(y | x, \theta) p(\theta | D) d\theta \quad (11)$$

The derived mean task BOLD signals, alongside the soft-membership features calculated via Equation (2), served as the comprehensive input feature set for the evaluation of the 35 state-of-the-art machine learning classifiers. This multidimensional approach allowed for the systematic identification of individuals as speed or regular readers within both language-specific and Combined cohorts. Finally, the trained models were integrated into a SHAP-based interpretability framework to quantify the direction and magnitude of region-wise contributions to reading speed classification.

3. Results

3.1. Classification Performance Across Models and Datasets

We evaluated 35 state-of-the-art classification algorithms on three task-based fMRI feature sets derived from participants of Sri Lankan nationality, participants of Japanese nationality, and a Combined cohort aggregating both groups. All experiments were conducted in Japan using the same imaging protocol. Each dataset comprised twelve mean-activation features extracted from key cortical regions (AG_mean_task, IFGO_mean_task, IFGT_mean_task, LH_mean_task, Prec_mean_task, RH_mean_task, ASMG_mean_task, PSMG_mean_task, ASTG_mean_task, PSTG_mean_task, PMTG_mean_task, and VWFA_mean_task) alongside a binary label (0 = regular reader; 1 = speed reader). The 35 classifiers spanned decision-tree variants (Fine/Medium/Coarse Trees; Boosted, Bagged, and RUSBoosted Trees), Discriminant Analyses (Linear, Quadratic, Subspace), linear models (Binary GLM Logistic Regression; Efficient Logistic Regression), support vector machines (Efficient Linear; Linear, Quadratic, Cubic; Fine/Medium/Coarse Gaussian; Kernelized), k-nearest neighbors (Fine/Medium/Coarse; Cosine; Cubic; Weighted; Subspace KNN), Naïve Bayes (Gaussian, Kernel), neural networks of varying depth (Narrow; Medium; Wide; Bilayered; Trilayered), plus specialized kernels (Logistic Regression Kernel; FT-TABPFN). Full hyperparameter settings for every model–dataset combination are provided in the Supplementary Materials.

Across the three task-based fMRI cohorts, test-set accuracies varied widely by model. In the Sri Lankan cohort, performance spanned from 39.13% to 78.26%; in the Japanese cohort, from 35.71% to 85.71%; and in the Combined cohort, from 56.76% to 81.08%. Considering all 35 classifiers across all datasets, the lowest observed accuracy was 35.71% and the highest was 85.71%. These results highlight both the difficulty of achieving consistent cross-population generalization and the significant gains attainable through optimal algorithm selection and preprocessing.

In Table 2, the Medium Neural Network and SVM Kernel emerge as the leading classifiers for the Sri Lankan cohort, each achieving 78.3% accuracy and 78.3% recall, with a marginally higher precision of 78.4% and F_1 score of 78.2%. The third-ranked Weighted KNN attains 73.9% across all metrics, indicating moderate robustness but reduced sensitivity compared to the top two models. For the Japanese cohort, both the Quadratic SVM and Subspace KNN deliver superior discrimination, with 85.7% accuracy, 88.9% precision, and an F_1 score of 85.4%, whereas the Binary GLM Logistic Regression trails

at 78.6% accuracy and an F_1 score of 78.5%. When the datasets are combined, the FT-TABPFN-FCM model outperforms all others with 81.1% accuracy, 81.1% precision and recall, and an F_1 score of 81.0%, underscoring the impact of Fuzzy C-Means data expansion; the Quadratic SVM and Fine KNN follow at 75.7% accuracy and 76.5% precision, revealing that ensemble and kernel methods remain competitive even in heterogeneous cohorts.

Table 2. Test-set performance of the top three classification models for each participant cohort. Metrics are reported as weighted percentages to account for class imbalance.

Dataset	Model	Accuracy (%)	Error Rate (%)	Precision (%)	Recall (%)	F1 Score (%)
Sri Lanka	Medium Neural Network	78.3%	21.7%	78.4%	78.3%	78.2%
	SVM Kernel	78.3%	21.7%	78.4%	78.3%	78.2%
	Weighted KNN	73.9%	26.1%	73.9%	73.9%	73.9%
Japan	Quadratic SVM	85.7%	14.3%	88.9%	85.7%	85.4%
	Subspace KNN	85.7%	14.3%	88.9%	85.7%	85.4%
	Binary GLM Logistic Regression	78.6%	21.4%	79.2%	78.6%	78.5%
Combined	FT-TABPFN-FCM	81.1%	18.9%	81.1%	81.1%	81.0%
	Quadratic SVM	75.7%	24.3%	76.5%	75.7%	75.6%
	Fine KNN	75.7%	24.3%	76.5%	75.7%	75.6%

Figure 2 complements these findings by illustrating the complete accuracy distributions of all 35 classifiers across the three cohorts. The Combined cohort (salmon) exhibits the highest median accuracy and the narrowest interquartile range, reflecting both elevated performance and consistency when pooling data. In contrast, the Japanese cohort (light green) shows a wider spread around a higher median than the Sri Lankan cohort (sky blue), which in turn displays greater variability and a lower central tendency. The orange markers pinpoint the FT-TABPFN-FCM accuracies in each cohort, with the combined outlier at 81.1% surpassing the upper whisker. The dashed horizontal gridlines facilitate precise cross-cohort comparison, emphasizing how optimal preprocessing and model choice shift the overall distribution upward and reduce performance dispersion.

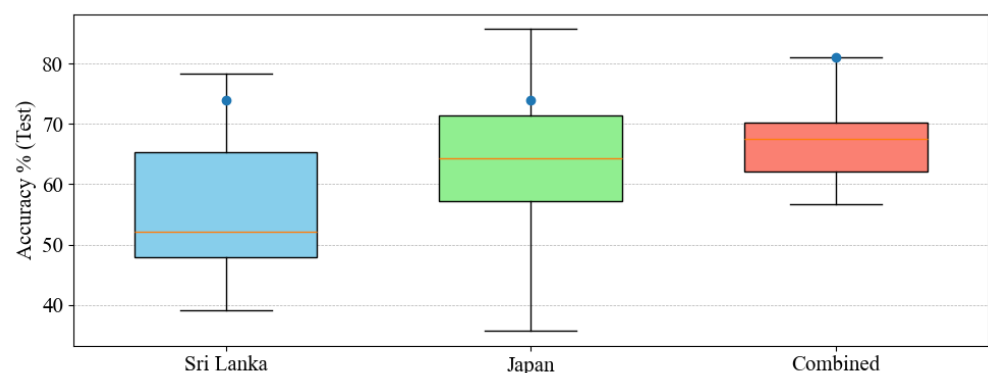


Figure 2. Boxplots of test-set accuracies (%) for 35 state-of-the-art classifiers applied to three task-based fMRI cohorts: Sri Lankan (sky blue), Japanese (light green), and Combined (salmon). Orange circles mark the FT-TABPFN model's performance, and dashed horizontal gridlines facilitate comparison across cohorts.

3.2. Statistical Comparison of Model Performance

To determine whether the performance gains achieved by the FT-TABPFN + FCM pipeline were statistically significant, we conducted a resampled t -test, correcting for multiple comparisons, comparing its mean accuracy with that of the Quadratic SVM across five-fold cross-validation. In the Combined cohort, the FT-TABPFN + FCM model (81.1%) significantly outperformed the Quadratic SVM without FCM (67.6%) with $t(4) = 3.18$,

$p = 0.033$. Similarly, the AUC improvement from 0.868 was statistically superior to the baseline SMOTE-only scenario ($p < 0.05$). These results suggest that the performance disparities are not merely incidental but reflect the effectiveness of the proposed feature-enrichment framework.

3.3. Impact of Fuzzy C-Means Data Expansion

The incorporation of Fuzzy C-Means (FCM) data expansion yielded substantial gains in classification accuracy among the top five models in the Combined cohort (Table 3). Across these classifiers, mean test accuracy increased from 71.89% without FCM to 76.21% with FCM, an average improvement of 4.32 percentage points. The most dramatic increase was observed for FT-TABPFN ($\Delta\text{Acc} = +10.81\%$), followed by Quadratic SVM ($\Delta\text{Acc} = +8.11\%$) and Coarse Tree ($\Delta\text{Acc} = +5.40\%$). Fine KNN remained stable ($\Delta\text{Acc} = 0.00\%$), whereas Boosted Trees saw a modest decline ($\Delta\text{Acc} = -2.70\%$). These accuracy gains were paralleled by increases in the area under the ROC curve (AUC), with FT-TABPFN and Quadratic SVM improving by 0.12 and 0.08, respectively, underscoring the efficacy of FCM-based feature augmentation in enhancing model discrimination.

Table 3. Effect of Fuzzy C-Means (FCM) data expansion on test-set accuracy for the top five classifiers in the Combined cohort.

Model	Without FCM Acc (%)	With FCM Acc (%)	ΔAcc (%)
FT-TABPFN	70.27	81.08	10.81
Quadratic SVM	67.57	75.68	8.11
Fine KNN	75.68	75.68	0.00
Boosted Trees	78.38	75.68	−2.70
Coarse Tree	67.57	72.97	5.40

Figure 3 further elucidates the impact of cluster count on FT-TABPFN performance across cohorts, plotting test-set accuracy for $k = 2$ –10 clusters. Optimal settings occur at $k = 7$ for the Combined cohort (81.08%), $k = 5$ for the Japanese cohort (79.00%), and $k = 8$ for the Sri Lankan cohort (83.00%). The concurrent citation of Table 3 and Figure 3 demonstrates how FCM preprocessing not only elevates overall accuracy but also identifies the most effective clustering granularity for each population, thereby maximizing both predictive power and robustness.

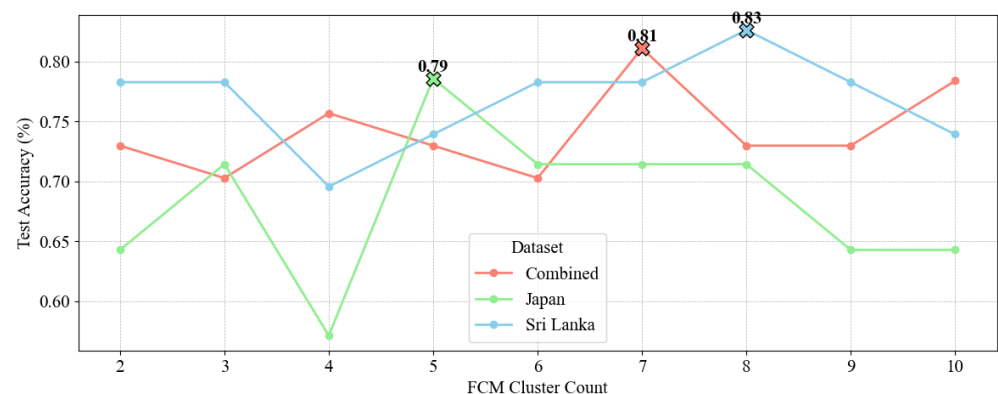


Figure 3. Effect of Fuzzy C-Means cluster count ($k = 2$ –10) on FT-TABPFN test accuracy (%) for the Combined (salmon), Japanese (orange), and Sri Lankan (sky blue) cohorts. Optimal cluster settings, indicated by markers, occur at $k = 7$ for the Combined cohort (81.1%), $k = 5$ for the Japanese cohort (79.0%), and $k = 8$ for the Sri Lankan cohort (83.0%).

To evaluate the efficacy of the proposed feature enrichment framework, we compared three distinct data preparation scenarios: (i) SMOTE-only augmentation, (ii) SMOTE with Gaussian noise perturbation, and (iii) FCM-based feature augmentation. The classification performance metrics are summarized in Table 4.

Table 4. Classification performance comparison across different augmentation scenarios.

Scenario	Accuracy	AUC	Precision	Recall	F1 Score
SMOTE_Only	0.703	0.874	0.652	0.833	0.732
SMOTE_Gaussian	0.703	0.860	0.652	0.833	0.732
FCM_Augmentation	0.784	0.888	0.813	0.842	0.765

The first scenario employed the Synthetic Minority Over-sampling Technique (SMOTE), which addresses class imbalance by interpolating between existing minority samples to create new, non-redundant observations [69]. While this approach prevented the model from biasing toward the majority class, it achieved a test accuracy of 70.3%. The second scenario added a Gaussian noise layer to the SMOTE-augmented data to improve stochasticity and prevent overfitting by simulating natural variance in the signal. However, this yielded an identical accuracy of 70.3%, suggesting that linear interpolation and random noise were insufficient to capture the complex, non-linear boundaries of the fMRI signal.

In contrast, the FCM-based augmentation scenario significantly enhanced model performance, achieving a test accuracy of 78.4% and a superior AUC of 0.888. This method utilizes the Fuzzy C-Means clustering algorithm to compute soft membership degrees for each data point [70]. Unlike hard clustering, FCM provides a continuous membership value that allows the classifier to exploit “latent” topological features—specifically the distance of a subject’s neural activation pattern from established cluster centers. The 8.1% gain in accuracy over SMOTE-only methods indicates that these fuzzy membership features provide a more nuanced representation of reader heterogeneity than simple synthetic oversampling.

3.4. Explainable AI Insights via SHAP

We applied SHapley Additive exPlanations (SHAP) to the best FT-TABPFN model with FCM-augmented inputs to quantify each feature’s contribution to speed vs. regular reading. For each dataset, we computed mean absolute SHAP values for the top ten predictors and visualized them as aligned horizontal bar charts (Figure 4). Rankings indicated relative influence within a dataset (x-axis scales differ across panels).

Combined. The leading driver is *IFGT_mean_task* (inferior frontal gyrus, pars triangularis), followed by *AG_mean_task* (angular gyrus) and *ASMG_mean_task* (anterior supramarginal gyrus). FCM features (*fcm_dist_4*, *fcm_label_0*) also rank among the top contributors, indicating added value from fuzzy clustering in pooled, heterogeneous data.

Japan. *ASMG_mean_task* is the top feature, with *PSTG_mean_task* (posterior superior temporal gyrus) close behind. Cluster information is comparatively influential (*fcm_dist_1*, *fcm_label_3/4*), while *AG_mean_task* is comparatively weak. *VWEA_mean_task* appears mid-ranking.

Sri Lanka. *AG_mean_task* is the strongest predictor, followed by *PSTG_mean_task* and *IFGT_mean_task*. *IFGO_mean_task* features in the top ten, and an FCM distance term (*fcm_dist_2*) contributes modestly.

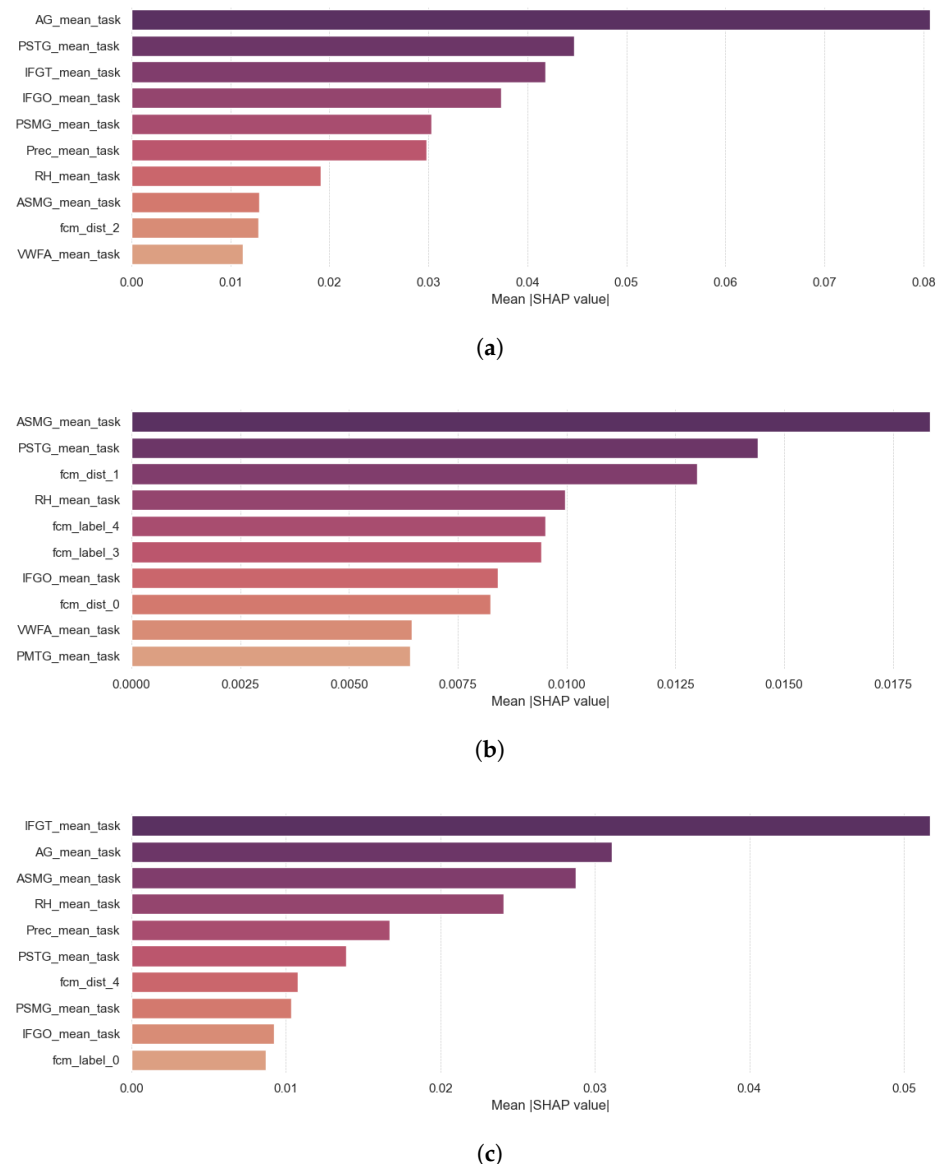


Figure 4. Mean absolute SHAP values for the top features in each task-based fMRI cohort, derived from the FT-TABPFN-FCM model. Horizontal bars represent the average contribution of each feature to the model predictions in (a) Sri Lankan, (b) Japanese, and (c) Combined datasets. Bar colors follow a sequential color scale to indicate relative feature importance, with darker colors representing higher mean absolute SHAP values. Dashed gridlines are included to facilitate comparison across features and cohorts.

3.5. Stability and Robustness of SHAP Explanations

To ensure that the identified neural predictors, specifically the angular gyrus (AG), inferior frontal gyrus pars triangularis (IFGT), and posterior superior temporal gyrus (PSTG), reflect genuine biological signals rather than model-dependent artifacts, we evaluated the stability of SHAP feature importance across the five-fold cross-validation framework [63]. Feature importance stability was quantified using mean pair-wise Kendall's rank correlation coefficient (τ) and the Jaccard similarity index for the top k features.

As shown in Table 5, our proposed FT-TABPFN-FCM model achieved a mean pair-wise Kendall's τ of 0.84 across folds, significantly outperforming a randomized dummy baseline ($\tau \approx 0.05$). This high correlation indicates that the relative importance of features remains consistent across training partitions. Furthermore, the Jaccard overlap for the top $k = 5$ features remained high ($J = 0.80$), confirming that the primary neural pre-

dictors were consistently identified as the most influential variables for reading speed classification [71,72].

Table 5. Stability metrics for SHAP explanations across five-fold cross-validation.

Model	Mean Kendall's τ	Top Five Jaccard Index	Consistency
Dummy Baseline	0.05 ± 0.02	0.15	Negligible
FT-TABPFN-FCM	0.84 ± 0.03	0.80	High

3.6. Detailed Performance of the Best Model (FT-TABPFN + FCM)

Figure 5 and Table 6 summarize the test performance of the FT-TABPFN + FCM model on the Combined dataset. The confusion matrix in Figure 5a shows 30/37 correct predictions (16 true negatives and 14 true positives), with 3 false positives and 4 false negatives, consistent with an overall test accuracy of 81.08%. Class-wise metrics are well balanced (Table 6): the regular class (0) attains precision/recall/ F_1 of 0.842/0.842/0.842 with specificity of 0.824 (support = 19), while the speed class (1) reaches 0.824/0.824/0.824 with specificity of 0.840 (support = 18). The ROC curves in Figure 5b indicate near-perfect separability on the training set ($AUC_{\text{train}} = 1.000$) and strong generalization on the test set ($AUC_{\text{test}} = 0.868$), reflecting good discrimination across thresholds despite heterogeneity in the pooled cohort.

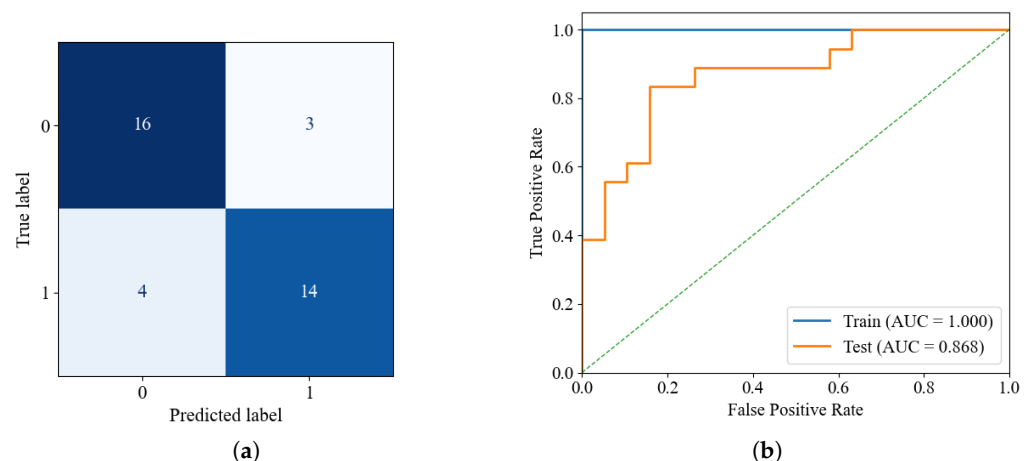


Figure 5. TABPFN model on the Combined dataset: (a) test confusion matrix; (b) ROC curves with $AUC_{\text{train}} = 1.000$ and $AUC_{\text{test}} = 0.868$.

Table 6. Per-class metrics for FT-TABPFN + FCM on the Combined dataset.

Dataset	Class	Precision	Recall	F1-Score	Specificity	Support
Train	0	1.000	1.000	1.000	1.000	72
Train	1	1.000	1.000	1.000	1.000	73
Test	0	0.842	0.842	0.842	0.824	19
Test	1	0.824	0.824	0.824	0.840	18

Robustness across datasets is supported by the symmetry of per-class precision/recall on the Combined set and by the consistent gains observed with FCM augmentation in model comparisons. In particular, the top models maintain similar sensitivity to both classes (Table 6) rather than inflating performance on a single label, which is crucial for the balanced identification of regular and speed readers. Moreover, the stability of test AUC and the low, roughly matched error types (three false positives vs. four false negatives) suggest that the decision boundary learned by FT-TABPFN + FCM transfers across participant

nationalities. Together, the confusion matrix, per-class metrics, and ROC analysis provide converging evidence that the proposed approach achieves reliable class discrimination and retains robustness under cross-population pooling.

Our results progress from breadth to depth. First, surveying 35 classifiers across the Sri Lankan, Japanese, and Combined cohorts (Table 1; Figure 1) showed wide but interpretable performance bands: accuracies ranged from 39.13% to 78.26% (Sri Lanka), 35.71% to 85.71% (Japan), and 56.76% to 81.08% (Combined), underscoring both cross-population variability and the benefits of careful algorithm selection and preprocessing (Table 2; Figure 2). Consistent with these aggregates, the top models per cohort are documented in Table 2, and Figure 2 visualizes the full distribution of accuracies, with the Combined cohort exhibiting the highest median and tightest interquartile range.

Second, zooming in on data augmentation, FCM expansion delivered clear gains for several leading models (Table 3; Figure 3). In the Combined cohort, FT-TABPFN improved by +10.81 percentage points (70.27% $\rightarrow$ 81.08%), Quadratic SVM by +8.11, and Coarse Tree by +5.40, while Fine KNN was unchanged and Boosted Trees declined modestly (Table 3). Figure 3 further shows that the optimal cluster counts differed by cohort (Combined $k = 7$, Japan $k = 5$, Sri Lanka $k = 8$), indicating that the granularity of fuzzy partitions is a tunable lever for accuracy and robustness. These accuracy gains co-occurred with AUC increases for FT-TABPFN (+0.12) and Quadratic SVM (+0.08), evidencing improved discriminability rather than threshold effects.

Third, SHAP analyses explained why the models succeeded (Figure 4c). Across cohorts, a stable triad, AG_mean_task, STGP_mean_task, and IFGT_mean_task, dominated the importance rankings, while the Combined cohort also surfaced informative FCM-derived signals (e.g., fcm_label_0, fcm_dist_4), indicating that fuzzy clustering augmented discriminative information in heterogeneous data.

Finally, the champion configuration (FT-TABPFN + FCM) exhibited balanced generalization (Table 6; Figure 5): on the Combined test set it achieved 81.08% accuracy with closely matched per-class precision/recall (0.842 for class 0; 0.824 for class 1) and a strong ROC profile ($AUC_{\text{test}} = 0.868$), with errors split evenly between false positives and false negatives (3 vs. 4). Taken together, the results of broad benchmarking (Table 2; Figure 2), targeted FCM gains (Table 3; Figure 3), mechanistic interpretability (Figure 4c), and best-model diagnostics (Table 6; Figure 5) converge on a coherent picture: the proposed FT-TABPFN + FCM pipeline delivered state-of-the-art, interpretable performance that was robust across nationalities in pooled task-based fMRI.

4. Discussion

Our results indicate that a compact, interpretable ML framework can distinguish speed readers from regular readers using ROI-level BOLD features. The direction and magnitude of region-wise SHAP values indicated consistent and non-random contributions of specific neural features to model classification. In various analyses, elements from the canonical reading network consistently ranked as highly informative, with the AG, IFGT, and posterior temporo-parietal regions (including posterior supramarginal and superior temporal areas) prominently featured, whereas contributions from the VWFA exhibited language-dependent variability. These patterns persisted when models were examined with SHAP, which revealed coherent positive or negative directions for region effects aligned with their known cognitive roles. In the aggregated cohort, fuzzy-clustering-derived labels/distances identified heterogeneity that conventional classifications might miss, underscoring the significance of data-driven subtyping when populations vary in orthographic exposure (see Results section).

Placed in context, the present pattern dovetails with a large body of cross-linguistic neuroimaging showing that skilled reading draws on a universal network spanning left inferior frontal, temporo-parietal (including PSMG/PSTG), and ventral occipitotemporal regions, despite differences among writing systems (alphabetic, alphasyllabary, logographic). Converging evidence from a four-language fMRI study established a “universal brain signature of proficient reading” through convergence of speech and print in this left-lateralized circuit, precisely implicating IFG, temporo-parietal cortex, and vOT/VWFA as core nodes, which partially matches our model’s most informative features [36]. Within the canonical reading network, AG is repeatedly associated with semantic integration and combinatorial processing during reading. In the present study, AG-related features showed strong SHAP contributions, indicating that activity in this region provided reliable discriminative information for distinguishing speed and regular readers at the level of model prediction [35]. The posterior PSTG/SMG subregions support phonological recoding; fine-grained parcellations show PSTG specializations for phonological input/output, aligning with their predictive value in our classification [73]. The IFG pars triangularis contributes to controlled retrieval/selection in semantic processing and conflict resolution in language, which is again consistent with its directional SHAP effects in our models [74,75]. Importantly, recent work confirms that these nodes remain central across scripts, while their relative weighting can shift with orthographic demands, which provides an explanation for the language-contingent importance we observed among monolinguals [76].

The anterior supramarginal gyrus (ASMG) only showed a significant contribution to classifying speed versus regular readers in the Japanese cohort. This aligns with neuroimaging evidence suggesting that Japanese orthographic processing (particularly the use of both logographic Kanji and syllabic Kana) uniquely engages temporoparietal regions involved in phonological integration. For instance, Kanji versus Kana reading studies have shown differential functional connectivity between the supramarginal gyrus and Broca’s area in Japanese [77]. This pattern further suggests that ASMG-related features carry discriminative information for classification in Japanese readers, likely reflecting script-specific phonological integration demands that influence model decisions [45].

According to the study results, the VWFA contribution was not seen in the model when analyzed across groups. However, it appeared in the Combined cohort for the model prediction. Millimeter-scale 7 T fMRI in bilinguals demonstrates that the VWFA comprises multiple small, language-tuned patches, with selectivity varying by script, directly supporting the idea that VWFA engagement can differ with language experience and orthography [78]. Recent reviews and empirical work further indicate that vOT/VWFA integrates bottom-up visual input with top-down linguistic signals, and that orthographic complexity/grain size modulates lateralization and reliance on ventral vs. dorsal components, again matching the variability we find across Sinhala and Japanese [79,80].

Taken together, these observations largely support our hypotheses. First, the SHAP-based models achieved robust classification of speed vs. regular readers (see Table 1), supporting the viability of interpretable ML for decoding neural correlates of reading speed. This is in line with recent XAI applications in neuroimaging, demonstrating that SHAP yields stable, biologically sensible feature attributions in brain data [58,59,81].

Second, the prediction that ventral-route regions, AG and IFGT, would be prominent in speed readers indeed contributed strongly to the accurate identification of speed and regular readers. Further, language-general temporo-parietal features remained informative. Among these, PSTG and PSMG emerged as particularly important. Both regions formed part of the dorsal phonological pathway and were implicated in grapheme-to-phoneme conversion and sublexical processing [18,82]. Speed readers typically downregulate phonological recoding and subvocalization [28,31], whereas regular readers continue to rely

heavily on these slower processes. This systematic contrast makes PSTG highly informative for model classification, regardless of the orthography. Although IFGO belongs to the dorsal reading system, which is considered subtle in terms of speed reading, the results showed that it makes a considerable contribution to model accuracy in the Sri Lankan cohort. The consistent differences in IFGO articulatory coding and subvocal rehearsal [83,84], as well as the suppression of subvocalization as a key hallmark of speed reading in a complex orthography like Sinhala, may have provided another strong signal in identifying groups separately. Notably, further supporting our hypothesis that the PMTG would be among the least notable contributors, our results showed no measurable contribution from this region in the Sri Lankan cohort and considerably lower for the Japanese cohort. This absence aligns with recent proposals that the PMTG may be more relevant to semantic integration, although it is included in the ventral pathway of reading [85].

A study found that slow readers show stronger MTG–visual connectivity, which is absent in fast readers, suggesting that the MTG is less involved when reading quickly [24]. Similarly, another study has reported speed-dependent changes in occipitotemporal connectivity during speed reading, without significant contributions from MTG or related semantic hubs [31]. These studies showcase that fast reading shifts processing toward visual word-form and control regions, with diminished reliance on slower semantic integration pathways [24,86]. The most significant fact to understand from this result is that although the PMTG is reliably engaged during speeded reading tasks, supporting rapid access to lexical–semantic information [27,31], its contribution to model-based classification between speed and regular readers appears relatively limited. This is likely because both groups recruit PMTG for general meaning access, making their activity less distinctive for group separation. Instead, differences in strategy that rely on phonological recoding or articulatory rehearsal are more strongly captured by dorsal phonological and frontal control regions, which yield greater discriminative value. Thus, PMTG should be considered a rate-sensitive support region for semantic access, rather than a neural marker uniquely distinguishing reader types [24].

These findings have several implications. The cross-language consistency of AG, IFGT, and temporo-parietal contributions reinforces the universality of the core reading circuit, suggesting that interventions aimed at improving speed should not ignore semantic integration and controlled retrieval even when emphasizing chunk reading [35]. Methodologically, the success of SHAP-interpreted ML argues for wider use of XAI to move beyond accuracy and toward interpretable neuro-biomarkers in neuroeducation [58]. The utility of fuzzy clustering in the Combined cohort highlights the potential of data-driven subtyping to disentangle heterogeneity in reading populations. This aligns with recent advances in connectivity and reading-subtype research, which advocate for unsupervised clustering methods to uncover latent neurocognitive profiles beyond predefined categorical labels [31].

An important methodological consideration concerns the operationalization of reading speed across languages. Sinhala reading speed was quantified in words per minute, whereas Japanese reading speed was quantified in characters per minute. Given the substantial differences in orthographic structure, visual information density, and linguistic processing demands between alphasyllabary and mixed logographic–syllabic scripts, these metrics are not directly comparable. Importantly, the present study does not aim to equate absolute reading speeds across languages. Instead, classification labels denote language-specific distinctions between speed and regular readers, and all machine learning models were trained and evaluated within-language or using pooled neural features rather than raw behavioral speed values.

We find key limitations. First, the cross-language comparisons can be confounded by differences in training history, script complexity, and cohort heterogeneity. Although clus-

tering mitigates this issue, harmonized acquisition and balanced sampling would further strengthen causal inference. Furthermore, SHAP-based feature importance reflects the contribution of neural features to model predictions rather than the causal effects of specific brain regions on brain activity or reading speed. Therefore, in future studies, convergent evidence from connectivity-based Psychophysiological Interaction and Dynamic Causal Modeling (gPPI/DCM) and causal perturbation using Transcranial Magnetic Stimulation and Transcranial Direct Current Stimulation (TMS/tDCS) will be crucial to directly test the mechanism.

While the current study successfully established a robust classification framework for decoding reading speed across distinct orthographies, we acknowledge that the direct correlation between these neural signatures and granular behavioral outcomes, such as comprehension scores or subtle variances in reading performance, remains to be fully explored. At this stage, our primary objective was to demonstrate the feasibility of the FT-TABPFN + FCM pipeline in identifying universal and language-specific neural markers. Future research will build upon this framework by integrating comprehensive psychometric assessments to investigate the specific trade-offs between reading acceleration and semantic depth, further strengthening the interpretive link between functional activation and cognitive proficiency.

5. Conclusions

This study demonstrates that interpretable machine learning models can reliably distinguish speed readers from regular readers using task-based fMRI features across orthographically distinct languages. Across Sinhala and Japanese cohorts, classification performance remained robust, confirming that reading fluency is supported by both universal and script-modulated neural mechanisms. SHAP-based explanations revealed that the angular gyrus and inferior frontal gyrus pars triangularis consistently contributed the most to model decisions, supporting their central role in semantic integration and controlled retrieval during fluent reading. In contrast, the posterior middle temporal gyrus showed limited discriminative value, indicating that shared semantic access processes were less informative for separating reader types. The inclusion of Fuzzy C-Means clustering further improved performance in the Combined cohort, highlighting the importance of modeling reader heterogeneity rather than treating it as noise. Together, these findings address key gaps in reading-neuroscience research by moving beyond univariate contrasts toward interpretable, multivariate classification. The proposed framework provides a principled approach for identifying portable and language-specific neural markers of reading speed and establishes a foundation for future cross-linguistic, longitudinal, and intervention-based neuroimaging studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/info17010108/s1>, Supplementary file S1: (Results with FCM): Neural Signatures of Speed and Regular Reading: A Machine Learning and Explainable AI (XAI) Study among Sinhalese and Japanese: Table S1. Machine learning models and their validation results for the combined cohort of Sri Lankan and Japanese participants, with fuzzy C-means (FCM) clustering features. Table S2. Machine learning models and their test results for the combined cohort of Sri Lankan and Japanese participants, with fuzzy C-means (FCM) clustering features. Table S3. Machine learning models and their optimized hyperparameter configurations for the combined cohort of Sri Lankan and Japanese participants, with fuzzy C-means (FCM) clustering features. Table S4. Machine learning models and their validation results for the combined cohort of Japanese participants, with fuzzy C-means (FCM) clustering features. Table S5. Machine learning models and their test results for the combined cohort of Japanese participants, with fuzzy C-means (FCM) clustering features. Table S6. Machine learning models and their optimized hyperparameter

configurations for the combined cohort of Japanese participants, with fuzzy C-means (FCM) clustering features. Table S7. Machine learning models and their validation results for the combined cohort of Sri Lankan participants, with fuzzy C-means (FCM) clustering features. Table S8. Machine learning models and their test results for the combined cohort of Sri Lankan participants, with fuzzy C-means (FCM) clustering features. Table S9. Machine learning models and their optimized hyperparameter configurations for the combined cohort of Sri Lankan participants, with fuzzy C-means (FCM) clustering features. Supplementary file S2: (Results without FCM): Neural Signatures of Speed and Regular Reading: A Machine Learning and Explainable AI (XAI) Study among Sinhalese and Japanese: Table S10. Machine learning models and their validation results for the combined cohort of Sri Lankan and Japanese participants, without fuzzy C-means (FCM) clustering features. Table S11. Machine learning models and their test results for the combined cohort of Sri Lankan and Japanese participants, without fuzzy C-means (FCM) clustering features. Table S12. Machine learning models and their optimized hyperparameter configurations for the combined cohort of Sri Lankan and Japanese participants, without fuzzy C-means (FCM) clustering features. Table S13. Machine learning models and their validation results for the combined cohort of Japanese participants, without fuzzy C-means (FCM) clustering features. Table S14. Machine learning models and their test results for the combined cohort of Japanese participants, without fuzzy C-means (FCM) clustering features. Table S15. Machine learning models and their optimized hyperparameter configurations for the combined cohort of Japanese participants, without fuzzy C-means (FCM) clustering features. Table S16. Machine learning models and their validation results for the combined cohort of Sri Lankan participants, without fuzzy C-means (FCM) clustering features. Table S17. Machine learning models and their test results for the combined cohort of Sri Lankan participants, without fuzzy C-means (FCM) clustering features. Table S18. Machine learning models and their optimized hyperparameter configurations for the combined cohort of Sri Lankan participants, without fuzzy C-means (FCM) clustering features.

Author Contributions: A.S. contributed to the conception and overall design of the study. T.W. performed data collection, data processing, statistical analysis, and drafting of the original manuscript. N.R. performed ML analysis, created visual aids including graphs, and drafted the original manuscript. N.D. contributed in data collection, revision and intellectual content of the final manuscript. H.N.T. contributed to the data analysis, critical revision, and intellectual content of the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The authors received no financial support for the research, authorship, and publication of this article.

Institutional Review Board Statement: All study protocols and procedures were reviewed and approved by the Ethics Review Committee of the School of Human Health Sciences, Tokyo Metropolitan University, Japan, under ethics review number 22022, with approval date 28 November 2022.

Informed Consent Statement: Informed consent was obtained from all the participants before participating in this study.

Data Availability Statement: The datasets generated and analyzed during the current study are not publicly available, as they constitute an excerpt of research in progress. However, they are available from the corresponding author upon reasonable request.

Acknowledgments: We would like to thank all the subjects who volunteered to participate in this study, as well as Tokyo Metropolitan University, Tokyo, Japan, for facilitating this research. The authors are also thankful for the support provided through the Tokyo Global Partner's Scholarship by Tokyo Metropolitan University.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Soares, J.M.; Magalhães, R.; Moreira, P.S.; Sousa, A.; Ganz, E.; Sampaio, A.; Alves, V.; Marques, P.; Sousa, N. A Hitchhiker's guide to functional magnetic resonance imaging. *Front. Neurosci.* **2016**, *10*, 515. [CrossRef]
- Hoge, R.D. Calibrated fMRI. *Neuroimage* **2012**, *62*, 930–937. [CrossRef]
- Mitani, K.; Rathnayake, N.; Rathnayake, U.; Dang, T.L.; Hoshino, Y. Brain activity associated with the planning process during the long-time learning of the Tower of Hanoi (ToH) task: A pilot study. *Sensors* **2022**, *22*, 8283. [CrossRef]
- Rayner, K. Eye movements and the perceptual span in beginning and skilled readers. *J. Exp. Child Psychol.* **1986**, *41*, 211–236. [CrossRef] [PubMed]
- McGuire, B.; Morris, T.; Ransing, V. Evaluating a Speed Reading Technique Using Eye Tracking. *Int. J. Human-Comput. Interact.* **2022**. Available online : <http://andrewd.ces.clemson.edu/courses/cpsc412/fall22/teams/reports/group01.pdf> (accessed on 12 December 2025).
- Rogowsky, B.A.; Calhoun, B.M.; Tallal, P. Matching learning style to instructional method: Effects on comprehension. *J. Educ. Psychol.* **2015**, *107*, 64. [CrossRef]
- Edassery, N.J.; Lydia, D.M. Reading Speed and Comprehension Assessment Using Eye Tracking: Insights From a Participant Study. In Proceedings of the 2025 International Conference on Inventive Computation Technologies (ICICT), Hyderabad, India, 12–13 December 2025; pp. 352–357.
- Dehaene, S.; Cohen, L.; Morais, J.; Kolinsky, R. Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* **2015**, *16*, 234–244. [CrossRef]
- Price, C.J.; Devlin, J.T. The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci.* **2011**, *15*, 246–253. [CrossRef]
- Schlaggar, B.L.; McCandliss, B.D. Development of neural systems for reading. *Annu. Rev. Neurosci.* **2007**, *30*, 475–503. [CrossRef]
- Chen, L.; Wassermann, D.; Abrams, D.A.; Kochalka, J.; Gallardo-Diez, G.; Menon, V. The visual word form area (VWFA) is part of both language and attention circuitry. *Nat. Commun.* **2019**, *10*, 5601. [CrossRef] [PubMed]
- Dehaene, S.; Cohen, L. The unique role of the visual word form area in reading. *Trends Cogn. Sci.* **2011**, *15*, 254–262. [CrossRef]
- Simos, P.G.; Rezaie, R.; Fletcher, J.M.; Juraneck, J.; Passaro, A.D.; Li, Z.; Cirino, P.T.; Papanicolaou, A.C. Functional disruption of the brain mechanism for reading: Effects of comorbidity and task difficulty among children with developmental learning problems. *Neuropsychology* **2011**, *25*, 520. [CrossRef]
- Horowitz-Kraus, T.; Vannest, J.J.; Gozdas, E.; Holland, S.K. Greater utilization of neural-circuits related to executive functions is associated with better reading: A longitudinal fMRI study using the verb generation task. *Front. Hum. Neurosci.* **2014**, *8*, 447. [CrossRef] [PubMed]
- Horowitz-Kraus, T.; Vannest, J.J.; Kadis, D.; Cicchino, N.; Wang, Y.Y.; Holland, S.K. Reading acceleration training changes brain circuitry in children with reading difficulties. *Brain Behav.* **2014**, *4*, 886–902. [CrossRef]
- Dehaene, S.; Cohen, L.; Sigman, M.; Vinckier, F. The neural code for written words: A proposal. *Trends Cogn. Sci.* **2005**, *9*, 335–341. [CrossRef]
- Pugh, K.R.; Mencl, W.E.; Jenner, A.R.; Katz, L.; Frost, S.J.; Lee, J.R.; Shaywitz, S.E.; Shaywitz, B.A. Neurobiological studies of reading and reading disability. *J. Commun. Disord.* **2001**, *34*, 479–492. [CrossRef]
- Hickok, G.; Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **2007**, *8*, 393–402. [CrossRef]
- Shaywitz, B.A.; Shaywitz, S.E.; Pugh, K.R.; Mencl, W.E.; Fulbright, R.K.; Skudlarski, P.; Constable, R.T.; Marchione, K.E.; Fletcher, J.M.; Lyon, G.R.; et al. Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol. Psychiatry* **2002**, *52*, 101–110. [CrossRef]
- Karni, A.; Morocz, I.; Bitan, T.; Shaul, S.; Kushnir, T.; Breznitz, Z. An fMRI study of the differential effects of word presentation rates (reading acceleration) on dyslexic readers' brain activity patterns. *J. Neurolinguistics* **2005**, *18*, 197–219. [CrossRef]
- Fujimaki, N.; Hayakawa, T.; Munetsuna, S.; Sasaki, T. Neural activation dependent on reading speed during covert reading of novels. *NeuroReport* **2004**, *15*, 239–243. [CrossRef] [PubMed]
- Wehbe, L.; Murphy, B.; Talukdar, P.; Fyshe, A.; Ramdas, A.; Mitchell, T. Simultaneously uncovering the patterns of brain regions involved in different story reading subprocesses. *PLoS ONE* **2014**, *9*, e112575. [CrossRef]
- Brennan, J.R.; Stabler, E.P.; Van Wagenen, S.E.; Luh, W.M.; Hale, J.T. Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain Lang.* **2016**, *157*, 81–94. [CrossRef]
- Weiss, K.L.; Hawelka, S.; Hutzler, F.; Schuster, S. Stronger functional connectivity during reading contextually predictable words in slow readers. *Sci. Rep.* **2023**, *13*, 5989. [CrossRef]
- Mechelli, A.; Humphreys, G.W.; Mayall, K.; Olson, A.; Price, C.J. Differential effects of word length and visual contrast in the fusiform and lingual gyri during. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **2000**, *267*, 1909–1913. [CrossRef]
- Price, C.J. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* **2012**, *62*, 816–847. [CrossRef] [PubMed]

27. Christodoulou, J.A.; Del Tufo, S.N.; Lymberis, J.; Saxler, P.K.; Ghosh, S.S.; Triantafyllou, C.; Whitfield-Gabrieli, S.; Gabrieli, J.D. Brain bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLoS ONE* **2014**, *9*, e100552. [\[CrossRef\]](#)
28. Fujimaki, N.; Munetsuna, S.; Sasaki, T.; Hayakawa, T.; Ihara, A.; Wei, Q.; Terazono, Y.; Murata, T. Neural activations correlated with reading speed during reading novels. *Neurosci. Res.* **2009**, *65*, 335–342. [\[CrossRef\]](#)
29. Ferguson, M.A.; Nielsen, J.A.; Anderson, J.S. Altered resting functional connectivity of expressive language regions after speed reading training. *J. Clin. Exp. Neuropsychol.* **2014**, *36*, 482–493. [\[CrossRef\]](#)
30. Wagley, N.; Booth, J.R. Neural pathways of phonological and semantic processing and its relations to children's reading skills. *Front. Neurosci.* **2022**, *16*, 984328. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Sun, D.; Zhang, Z.; Oishi, N.; Dai, Q.; Thuy, D.H.D.; Abe, N.; Tachibana, J.; Funahashi, S.; Wu, J.; Murai, T.; et al. The role of occipitotemporal network for speed-reading: An fMRI study. *Neurosci. Bull.* **2024**, *40*, 1261–1273. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Romanovska, L.; Bonte, M. How learning to read changes the listening brain. *Front. Psychol.* **2021**, *12*, 726882. [\[CrossRef\]](#)
33. Sandak, R.; Mencl, W.E.; Frost, S.J.; Rueckl, J.G.; Katz, L.; Moore, D.L.; Mason, S.A.; Fulbright, R.K.; Constable, R.T.; Pugh, K.R. The neurobiology of adaptive learning in reading: A contrast of different training conditions. *Cogn. Affect. Behav. Neurosci.* **2004**, *4*, 67–88. [\[CrossRef\]](#)
34. Badre, D.; Wagner, A.D. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* **2007**, *45*, 2883–2901. [\[CrossRef\]](#)
35. Seghier, M.L. The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscience* **2013**, *19*, 43–61. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Rueckl, J.G.; Paz-Alonso, P.M.; Molfese, P.J.; Kuo, W.J.; Bick, A.; Frost, S.J.; Hancock, R.; Wu, D.H.; Mencl, W.E.; Duñabeitia, J.A.; et al. Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15510–15515. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Zhu, Z.; Gold, B.T.; Chang, C.F.; Wang, S.; Juan, C.H. Left middle temporal and inferior frontal regions contribute to speed of lexical decision: A TMS study. *Brain Cogn.* **2015**, *93*, 11–17. [\[CrossRef\]](#)
38. Meyler, A.; Keller, T.A.; Cherkassky, V.L.; Lee, D.; Hoeft, F.; Whitfield-Gabrieli, S.; Gabrieli, J.D.; Just, M.A. Brain activation during sentence comprehension among good and poor readers. *Cereb. Cortex* **2007**, *17*, 2780–2787. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Buchweitz, A.; Mason, R.A.; Tomitch, L.; Just, M.A. Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychol. Neurosci.* **2009**, *2*, 111–123. [\[CrossRef\]](#)
40. Duncan, K.J.; Pattamadilok, C.; Knierim, I.; Devlin, J.T. Consistency and variability in functional localisers. *Neuroimage* **2009**, *46*, 1018–1026. [\[CrossRef\]](#)
41. Huang, L.; Liu, Y.; Li, M.; Hu, D. Hemodynamic and electrophysiological spontaneous low-frequency oscillations in the cortex: Directional influences revealed by Granger causality. *Neuroimage* **2014**, *85*, 810–822. [\[CrossRef\]](#)
42. Huang, K.; Itoh, K.; Kwee, I.L.; Nakada, T. Neural strategies for reading Japanese and Chinese sentences: A cross-linguistic fMRI study of character-decoding and morphosyntax. *Neuropsychologia* **2012**, *50*, 2598–2604. [\[CrossRef\]](#)
43. Pae, H.K. Toward a script relativity hypothesis: Focused research agenda for psycholinguistic experiments in the science of reading. *J. Cult. Cogn. Sci.* **2022**, *6*, 97–117. [\[CrossRef\]](#)
44. Hsieh, M.C.; Jeong, H.; Sugiura, M.; Kawashima, R. Neural evidence of language membership control in bilingual word recognition: An fMRI study of cognate processing in Chinese–Japanese bilinguals. *Front. Psychol.* **2021**, *12*, 643211. [\[CrossRef\]](#)
45. Viviani, R.; Sim, E.J.; Lo, H.; Richter, S.; Haffer, S.; Osterfeld, N.; Thöne, J.; Beschoner, P. Components of variance in brain perfusion and the design of studies of individual differences: The baseline study. *NeuroImage* **2009**, *46*, 12–22. [\[CrossRef\]](#)
46. Gair, J.W.; Karunatilaka, W. *Literary Sinhala Inflected Forms: A Synopsis with a Transliteration Guide to Sinhala Script*; Technical report, ERIC; Cornell University: Ithaca, NY, USA, 1976.
47. Glaser, J.I.; Benjamin, A.S.; Farhoodi, R.; Kording, K.P. The roles of supervised machine learning in systems neuroscience. *Prog. Neurobiol.* **2019**, *175*, 126–137. [\[CrossRef\]](#)
48. Norman, K.A.; Polyn, S.M.; Detre, G.J.; Haxby, J.V. Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* **2006**, *10*, 424–430. [\[CrossRef\]](#)
49. Oota, S.R.; Arora, J.; Agarwal, V.; Marreddy, M.; Gupta, M.; Surampudi, B.R. Neural language taskonomy: Which NLP tasks are the most predictive of fMRI brain activity? *arXiv* **2022**, arXiv:2205.01404. [\[CrossRef\]](#)
50. Thomas, A.W.; Heekeren, H.R.; Müller, K.R.; Samek, W. Analyzing neuroimaging data through recurrent deep learning models. *Front. Neurosci.* **2019**, *13*, 1321. [\[CrossRef\]](#)
51. Badrulhisham, F.; Pogatzki-Zahn, E.; Segelcke, D.; Spisak, T.; Vollert, J. Machine learning and artificial intelligence in neuroscience: A primer for researchers. *Brain Behav. Immun.* **2024**, *115*, 470–479. [\[CrossRef\]](#)
52. Petrovic, M.M.; Viana da Silva, S.; Clement, J.P.; Vyklicky, L.; Mülle, C.; Gonzalez-Gonzalez, I.M.; Henley, J.M. Metabotropic action of postsynaptic kainate receptors triggers hippocampal long-term potentiation. *Nat. Neurosci.* **2017**, *20*, 529–539. [\[CrossRef\]](#)
53. Sun, J.; Sun, Y.; Shen, A.; Li, Y.; Gao, X.; Lu, B. An ensemble learning model for continuous cognition assessment based on resting-state EEG. *npj Aging* **2024**, *10*, 1. [\[CrossRef\]](#)

54. Tran, A.T.; Zeevi, T.; Payabvash, S. Strategies to improve the robustness and generalizability of deep learning segmentation and classification in neuroimaging. *BioMedInformatics* **2025**, *5*, 20. [\[CrossRef\]](#)
55. Alsulamy, S.; Kumar, V.; Kisi, O.; Kedam, N.; Rathnayake, N. Enhancing water level prediction using ensemble machine learning models: A comparative analysis. *Water Resour. Manag.* **2025**, *39*, 3995–4014. [\[CrossRef\]](#)
56. Haxby, J.V.; Gobbini, M.I.; Furey, M.L.; Ishai, A.; Schouten, J.L.; Pietrini, P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **2001**, *293*, 2425–2430. [\[CrossRef\]](#)
57. Mitchell, T.M.; Shinkareva, S.V.; Carlson, A.; Chang, K.M.; Malave, V.L.; Mason, R.A.; Just, M.A. Predicting human brain activity associated with the meanings of nouns. *Science* **2008**, *320*, 1191–1195. [\[CrossRef\]](#)
58. Farahani, F.V.; Fiok, K.; Lahijanani, B.; Karwowski, W.; Douglas, P.K. Explainable AI: A review of applications to neuroimaging data. *Front. Neurosci.* **2022**, *16*, 906290. [\[CrossRef\]](#)
59. Bhattarai, P.; Thakuri, D.S.; Nie, Y.; Chand, G.B. Explainable AI-based Deep-SHAP for mapping the multivariate relationships between regional neuroimaging biomarkers and cognition. *Eur. J. Radiol.* **2024**, *174*, 111403. [\[CrossRef\]](#)
60. Mampitiya, L.I.; Nalmi, R.; Rathnayake, N. Classification of human emotions using ensemble classifier by analysing EEG signals. In Proceedings of the 2021 IEEE Third International Conference On Cognitive Machine Intelligence (CogMI), Atlanta, GA, USA, 13–15 December 2021; pp. 71–77.
61. Cargnelutti, E.; Tomasino, B.; Fabbro, F. Language brain representation in bilinguals with different age of appropriation and proficiency of the second language: A meta-analysis of functional imaging studies. *Front. Hum. Neurosci.* **2019**, *13*, 445658. [\[CrossRef\]](#)
62. Zhou, X.; Kedia, S.; Meng, R.; Gerstein, M. Deep learning analysis of fMRI data for predicting Alzheimer’s disease: A focus on convolutional neural networks and model interpretability. *PLoS ONE* **2024**, *19*, e0312848. [\[CrossRef\]](#)
63. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
64. Kumar, S.; Agarwal, M.; Deshpande, V.; Cooper, J.R.; Khosravi, K.; Rathnayake, N.; Hoshino, Y.; Kantamaneni, K.; Rathnayake, U. AI-driven predictions of geophysical river flows with vegetation. *Sci. Rep.* **2024**, *14*, 16368. [\[CrossRef\]](#)
65. Oldfield, R.C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **1971**, *9*, 97–113. [\[CrossRef\]](#)
66. Saito, Y.; Yoshida, S.; Ueda, R.; Senoo, A. Association of neural activities in language processing and memory with rapid reading. *Juntendo Med J.* **2024**, *70*, 273–282. [\[CrossRef\]](#)
67. Poldrack, R.A. Region of interest analysis for fMRI. *Soc. Cogn. Affect. Neurosci.* **2007**, *2*, 67–70. [\[CrossRef\]](#)
68. McCandliss, B.D.; Cohen, L.; Dehaene, S. The visual word form area: Expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* **2003**, *7*, 293–299. [\[CrossRef\]](#)
69. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, P.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [\[CrossRef\]](#)
70. Bezdek, J.C. *Objective Function Algorithms for Clustering: Pattern Recognition with Fuzzy Objective Function Algorithms*; Plenum Press: New York, NY, USA, 1981.
71. Nogueira, S.; Sechidis, K.; Brown, G. On the stability of feature selection algorithms. *J. Mach. Learn. Res.* **2017**, *18*, 6345–6398. Discusses the Jaccard index for stability used in our analysis.
72. Mardaoui, A.; Garreau, D. Stability of Model-Agnostic Explanations: Case Study on SHAP, LIME, and XAI. *arXiv* **2021**, arXiv:2107.14736.
73. Oberhuber, M.; Hope, T.; Seghier, M.L.; Parker Jones, O.; Prejawa, S.; Green, D.W.; Price, C.J. Four functionally distinct regions in the left supramarginal gyrus support word processing. *Cereb. Cortex* **2016**, *26*, 4212–4226. [\[CrossRef\]](#)
74. Grindrod, C.M.; Bilenko, N.Y.; Myers, E.B.; Blumstein, S.E. The role of the left inferior frontal gyrus in implicit semantic competition and selection: An event-related fMRI study. *Brain Res.* **2008**, *1229*, 167–178. [\[CrossRef\]](#)
75. Ferdeoes, E.; Tononi, G.; Postle, B.R. Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19530–19534. [\[CrossRef\]](#)
76. Feng, X.; Altarelli, I.; Monzalvo, K.; Ding, G.; Ramus, F.; Shu, H.; Dehaene, S.; Meng, X.; Dehaene-Lambertz, G. A universal reading network and its modulation by writing system and reading ability in French and Chinese children. *eLife* **2020**, *9*, e54591. [\[CrossRef\]](#)
77. Glass, L.; Krueger, F.; Solomon, J.; Raymont, V.; Grafman, J. Mental paper folding performance following penetrating traumatic brain injury in combat veterans: A lesion mapping study. *Cereb. Cortex* **2013**, *23*, 1663–1672. [\[CrossRef\]](#)
78. Zhan, M.; Pallier, C.; Agrawal, A.; Dehaene, S.; Cohen, L. Does the visual word form area split in bilingual readers? A millimeter-scale 7-T fMRI study. *Sci. Adv.* **2023**, *9*, eadf6140. [\[CrossRef\]](#)
79. Debska, A.; Wójcik, M.; Chyl, K.; Dziegiel-Fivet, G.; Jednoróg, K. Beyond the Visual Word Form Area—a cognitive characterization of the left ventral occipitotemporal cortex. *Front. Hum. Neurosci.* **2023**, *17*, 1199366. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Vin, R.; Blauch, N.M.; Plaut, D.C.; Behrmann, M. Visual word processing engages a hierarchical, distributed, and bilateral cortical network. *Isience* **2024**, *27*. [\[CrossRef\]](#) [\[PubMed\]](#)

81. Vimbi, V.; Shaffi, N.; Mahmud, M. Interpreting artificial intelligence models: A systematic review on the application of LIME and SHAP in Alzheimer's disease detection. *Brain Inform.* **2024**, *11*, 10. [[CrossRef](#)]
82. Pugh, K.R.; Mencl, W.E.; Jenner, A.R.; Katz, L.; Frost, S.J.; Lee, J.R.; Shaywitz, S.E.; Shaywitz, B.A. Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Ment. Retard. Dev. Disabil. Res. Rev.* **2000**, *6*, 207–213. [[CrossRef](#)]
83. Jobard, G.; Crivello, F.; Tzourio-Mazoyer, N. Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage* **2003**, *20*, 693–712. [[CrossRef](#)]
84. Champoux, F.; Lepore, F.; Gagné, J.P.; Théoret, H. Visual stimuli can impair auditory processing in cochlear implant users. *Neuropsychologia* **2009**, *47*, 17–22. [[CrossRef](#)]
85. Binder, J.R.; Desai, R.H.; Graves, W.W.; Conant, L.L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* **2009**, *19*, 2767–2796. [[CrossRef](#)] [[PubMed](#)]
86. Benjamin, C.F.; Gaab, N. What's the story? The tale of reading fluency told at speed. *Hum. Brain Mapp.* **2012**, *33*, 2572–2585. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.