



AI-Enhanced Child Handwriting Analysis: A Framework for the Early Screening of Dyslexia and Dysgraphia

Sahana Rangasrinivasan^{1,2} · M. S. Sumi Suresh^{1,2} · Abbie Olszewski³ · Srirangaraj Setlur¹ · Bharat Jayaraman⁴ · Venu Govindaraju¹

Received: 13 July 2024 / Accepted: 22 March 2025
© The Author(s) 2025

Abstract

Dyslexia and dysgraphia are two specific learning disabilities (SLDs) that are prevalent among children. To minimize the negative impact these SLDs have on a child's academic and social-emotional development, it is crucial to identify dyslexia and dysgraphia at an early age, enabling timely and effective intervention. The first step in this process is screening, which helps determine if a child requires further instruction or a more in-depth assessment. Current screening tools are expensive, require additional administration time beyond regular classroom activities, and are designed to screen exclusively for one condition, not for both dyslexia and dysgraphia, which often share some common behavioral characteristics. Most dyslexia screeners focus on speech and oral tasks and exclude writing activities. However, analyzing children's writing samples for behavioral signs of dyslexia and dysgraphia can offer valuable insights into the screening process, which can be time-consuming. As a solution, we propose a co-designed framework for building artificial intelligence (AI) tools that could boost the efficiency of screening and aid practitioners such as speech-language pathologists (SLPs), occupational therapists, general educators, and special educators by simplifying their tasks. This paper reviews current screening methods employed by practitioners, the use of AI-based systems in identifying dyslexia and dysgraphia, and the handwriting datasets available to train such systems. The paper also outlines a framework for developing an AI-integrated screening tool that can identify writing-based behavioral indicators of dyslexia and dysgraphia in children's handwriting. This framework can be used in conjunction with current screening tools like the Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC). The paper also proposes a methodology for collecting children's offline and online handwriting samples to build a valuable dataset for developing AI solutions. The proposed framework and data collection methodology are co-designed with SLPs, occupational therapists (OTs), special educators, and general educators to ensure the tool can provide explainable, actionable information that would be invaluable in a practical setting.

Keywords Artificial intelligence · Child handwriting · Dysgraphia · Dyslexia · Early screening

Sahana Rangasrinivasan and M. S. Sumi Suresh have contributed equally to this work.

✉ Sahana Rangasrinivasan
srangeasr@buffalo.edu

M. S. Sumi Suresh
sumisure@buffalo.edu

Abbie Olszewski
aolszewski@unr.edu

Srirangaraj Setlur
setlur@buffalo.edu

Bharat Jayaraman
bharat.jayaraman@aiarchicago.org

Venu Govindaraju
govind@buffalo.edu

¹ Department of Computer Science and Engineering,
University at Buffalo, Buffalo, NY 14260, USA

² Department of Computer Science and Engineering, Amrita
School of Computing, Amrita Vishwa Vidyapeetham,
Amritapuri, Kollam, Kerala 690525, India

³ College of Education and Human Development, University
of Nevada-Reno, Reno, NV 89557, USA

⁴ Amrita Institute of Advanced Research, Elburn, IL 60119,
USA

Introduction

Proficient speech, oral, and written language skills are essential for a child's academic success and social-emotional development. Delays in acquiring these abilities can negatively impact educational performance, social interactions, and future vocational opportunities. Speech-language disabilities are a significant concern, affecting over half of the students receiving special education and related services under the Individuals with Disabilities Education Act (IDEA) [1]. Among these, specific learning disabilities (SLDs) are the most prevalent category, accounting for 32% of the population [2]. This underscores the critical need for early screening and intervention, particularly for children struggling with SLDs, to provide them with the support they need to thrive.

Dyslexia and dysgraphia are two prominent forms of SLDs that significantly impact language abilities. According to the International Dyslexia Association (IDA) [3], dyslexia is a neurodevelopmental disorder characterized by difficulties with accurate and/or fluent word recognition and poor spelling abilities. These challenges typically stem from a deficit in the phonological component of language. Hence, dyslexia encompasses both spoken and written language skills. For example, children with dyslexia often struggle with phonological and phonemic awareness, such as identifying the first sound of a spoken word. In writing, these deficits manifest as difficulty in decoding (reading) or encoding (spelling) words due to poor grapheme-phoneme correspondences, which rely heavily on oral phonological skills.

In contrast, dysgraphia is a neurodevelopmental disorder specifically related to written expression. It can result from deficits in fine motor skills, spatial perception, working memory, orthographic coding, or organization. Consequently, children with dysgraphia may exhibit illegible handwriting, poor spelling, or both. Unlike dyslexia, dysgraphia impacts only written language.

Despite their differences, dyslexia and dysgraphia share overlapping behavioral indicators in writing, as illustrated in Fig. 1. Both disorders can lead to challenges with letter formation, legibility, and spelling. Poor working memory associated with either condition may further impair writing, affecting organization, punctuation, and adherence to grammatical conventions.

As dyslexia is a spectrum disorder, its severity can range from mild to severe. Children with mild dyslexia may not exhibit noticeable spoken language difficulties but often encounter challenges as language demands increase, particularly in academic writing tasks. Examining children's written language is especially valuable, as it can help identify students who might otherwise be missed

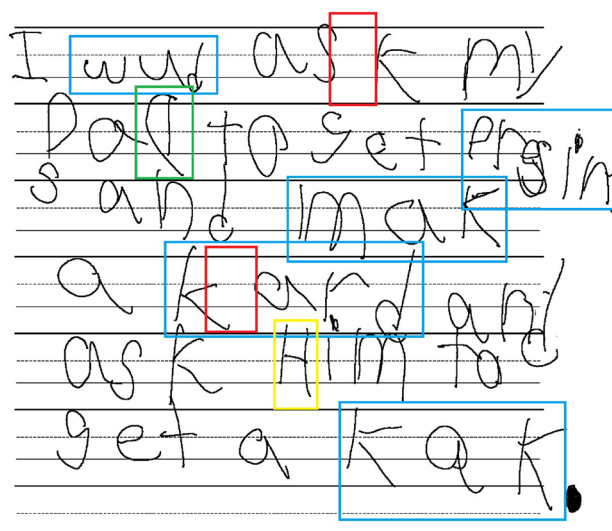


Fig. 1 Example of behaviors common to dyslexia and dysgraphia that can be found in a child's handwriting. Here, the colors of the bounding boxes represent different behaviors: *blue* represents spelling errors, *red* indicates atypical spacing between letters, *green* indicates reversed letters, and *yellow* represents inconsistent/lack of grammatical conventions

through spoken language screenings alone. Furthermore, since evaluating written language is already a part of daily academic activities, it provides a practical avenue for early detection and intervention. Thus, this paper focuses on the analysis of child handwriting to screen for dyslexia and dysgraphia.

Given the profound impact that dyslexia and dysgraphia have on children's academic and social development, early and accurate screening is crucial for timely intervention. However, traditional screening methods that focus on handwriting can be time-consuming, resource-intensive, and inconsistent, particularly when addressing the diverse manifestations of these disabilities. This highlights the need for innovative approaches to streamline the screening process while maintaining high accuracy and reliability.

The central thesis of this paper is that integrating artificial intelligence (AI) with screening boosts the efficiency of screening and aids practitioners such as speech-language pathologists (SLPs) by simplifying their tasks. Thus, this paper critically examines current screening methods to identify how AI can be applied to expedite screening for dyslexia and dysgraphia. Existing AI methods and child handwriting datasets designed for this task are also reviewed. Current datasets are limited in number and not publicly available. Moreover, the majority of datasets are geared toward dysgraphia identification and are not diverse enough to represent the wide range of handwriting variations. As a result, we present a methodology for collecting offline and online

child handwriting samples to develop a dataset that can be used for screening.

We also observe that literature on existing AI methods focuses on the use of machine learning (ML) models for mere diagnosis of dyslexia and dysgraphia, which may not be well-grounded or explainable. This hinders SLPs from providing ability-based interventions tailored to a child's unique strengths and challenges. Thus, we propose a framework to develop an AI-integrated screening tool that complements current screening methods and assists practitioners in identifying various writing-based behavioral indicators from children's handwriting samples. Specifically, this framework is discussed in conjunction with the recently developed Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC) [4], a screener that comprises 17 behavioral indicators associated with dysgraphia and dyslexia that can be observed during and after a writing task. The proposed framework will help scale the availability of SLPs for early screening and administration of ability-based interventions and provide opportunities to collaborate with occupational therapists (OTs), special education teachers, and general education teachers. To our knowledge, no other work has conducted such a comprehensive study to identify a framework to address this problem.

To ensure the development of effective and practical tools for screening, the proposed data collection methodology and framework are co-designed with practitioners such as SLPs, OTs, special educators, and general educators. By involving them in the design process, we can create tools that are not only scientifically sound but also tailored to the real-world needs of users. Their expertise helps in identifying key assessment criteria, ensuring the usability and accessibility of tools, and refining intervention strategies. Additionally, their participation in data collection ensures that the data gathered is comprehensive, relevant, and reflective of diverse educational contexts. Co-design fosters collaboration, aligns tools with educational goals, and leads to more effective support for children with dyslexia or dysgraphia.

The key contributions of this work are summarized as follows:

1. Review of current screening methods employed by practitioners, the use of AI-based systems in identifying dyslexia and dysgraphia, and the datasets available to train such systems.
2. Design of a framework for developing an AI-integrated screening tool that complements current screening methods for identifying writing-based behavioral indicators of dyslexia and dysgraphia.
3. Design of a method to collect children's offline and online handwriting samples needed to build the dataset underlying the AI-integrated screening tool.

4. Co-design of the proposed framework and data collection process with speech-language pathologists (SLPs), occupational therapists (OTs), special educators, and general educators.
5. Define specifications for the modules within the proposed framework to demonstrate how classifiers could be implemented for certain DDBIC indicators and to illustrate the challenges in developing the complex modules.

The remaining part of the paper is organized as follows: the “[Literature Review](#)” section discusses the works related to screening for dyslexia and dysgraphia, while the “[Methodology for Data Collection](#)” section provides an overview of the process for collecting the data to be used for screening. Thereafter, the “[Framework for Screening for Dyslexia and Dysgraphia](#)” section discusses the analytical modules that should be integrated in order to design the suite of AI tools for screening, and the “[Conclusion](#)” section presents our conclusions from this research.

Literature Review

This section reviews the existing literature on handwriting analysis to identify dysgraphia and dyslexia. We explore various manual and ML-based approaches, along with datasets that have been proposed to screen for these two SLDs efficiently. This review provides a foundation for understanding the state-of-the-art in handwriting analysis for screening and identifies gaps that our proposed framework aims to address.

Current Screening Methods Used by Practitioners

Screening is a crucial step in identifying behavioral characteristics associated with dyslexia and dysgraphia. It evaluates behaviors that indicate a need for further instruction, intervention, and potential evaluation [5]. This section reviews the screening tools currently used by practitioners to identify dyslexia or dysgraphia from a child's handwriting sample. Tables 1, 2, and 3 list some of the common tools used to screen for dyslexia, dysgraphia, or both, respectively.

Dyslexia Tools

Many states have legislation that requires screening children for dyslexia. Each state specifies the particular grades for screening and some states provide state-recognized screeners [6] such as AIMSweb Plus, DIBELS 8th Edition, Fastbridge, and PALS-K to name a few. These screeners may assess skills in the areas of phonological processing

Table 1 List of common tools used to screen for dyslexia in children

Name of screening tool	Grades/ages	Language	Administration	Length	Spelling	Computer	Classification accuracy
CUBED-3 [8]	Grades: Prek-3	English	Benchmark: beginning, middle, or end of year progress monitoring: every two weeks	5 min	Accuracy	Administration and scoring	Convincing
Lexia RAPID assessment [9]	Grades: K-12, Ages: 4–15 years	English	Benchmark: fall, winter, spring	Individual, small group, Large group—40 min	Accuracy	Administration and scoring	Convincing for 4th and 5th
PALS 1–3 [10]	Grades: 1–3	English	Benchmark: Fall, Spring	Individual, Small group, Large group—30 min	Accuracy	Administration and scoring	Partially convincing for 1st, 2nd, and 3rd

Table 2 List of common tools used to screen for dysgraphia in children using handwriting

Name of screening tool	Grades/ages	Language	Administration	Length	Targeted skill	Computer ^b
AimswbPlus spelling [12]	Grades: K-2	English	Benchmark: fall, winter, spring	Untimed	Spelling accuracy	N
Aimswb written expression [13]	Grades: 1–12	English	Benchmark: fall, winter, spring	–	Total words written, correct sequences written, words spelled correct	N
Beery-Buktenica developmental test of visual-motor integration (Beery VMI) [14]	Ages: 2–100	English	Test	10–15 min	Visual motor	Y
Detailed assessment of speed of handwriting (DASH) [15]	Grades: 4–9, Ages: 9–16 years	English	Test	5–30 min	Speed, legibility	N
Feifer assessment of writing (FAW) [16]	Grades: PreK-college Ages: 5–22	English	Test	15 min–1 h	Cognitive, motor, and linguistic processes	N
Lexercise dysgraphia test and online screener [17]	N/S ^a	English	Survey	5 min	Survey	Y
Movement assessment battery for children-2nd edition checklist (movement ABC) [18]	Ages: 5–12	English	Test	10 min	Checklist	N
Screener of handwriting proficiency [19]	Grades: K-5	English	Test	10–15 min	Formation, size, neatness, speed, posture, pencil grip, helper hand	Y
Alphabet writing fluency subtest of the Wechsler individual achievement test 4th edition (WIAT-4) [20]	Ages: 5–50 years 11 months	English	1 min	–	Letter writing fluency	Y

^aN/S, not specified^bY = Yes for computer involvement in administration and/or scoring, N = No for no computer involvement in administration and/or scoring

Table 3 A proposed tool used to screen for both dyslexia and dysgraphia in children using handwriting

Name of screening tool	Grade/age	Language	Task	Targeted skill
Dysgraphia and dyslexia behavioral checklist (DDBIC) [4]	Grades: K-5	English	Narrative or expository with pictorial cue	17 behaviors related to during writing, dysgraphia, dyslexia, or both

(rapid automatic naming, memory, awareness, phoneme-grapheme correspondence), word identification (speed and accuracy), nonword identification (speed and accuracy), listening comprehension, and reading comprehension [7]. Although a child's writing sample can provide valuable insights into the behavioral characteristics of dyslexia, few screeners evaluate a child's writing skills. Children with dyslexia may exhibit poor spelling, letter reversals, and atypical spacing in their writing. To our knowledge, only two dyslexia screeners globally address children's spelling: the Shaywitz Dyslexia Screener and the Test of Integrated Language and Literacy Skills Student Language Scale. These surveys have specific questions regarding spelling but do not analyze the writing itself. Three screeners directly examine spelling: CUBED-3, Lexia RAPID Assessment, and PALS 1–3. The Lexia RAPID Assessment and PALS 1–3 examine spelling accuracy while the CUBED-3 examines spelling, punctuation, and capitalization. While these screeners assess children's writing, they do not provide actionable insights on instruction, intervention, or additional supports, nor do they help differentiate these behavioral characteristics from dysgraphia.

Dysgraphia Tools

Unlike dyslexia, only a few states have legislation requiring screening for dysgraphia. For example, in 2017, Texas passed a bill mandating that schools screen for dyslexia and related disorders, including dysgraphia. Dysgraphia is considered a specific learning disability in written expression. Unfortunately, schools struggle to identify children with dysgraphia [11], which is further exacerbated by the limited dysgraphia screeners. Instead, dysgraphia is identified when testing for specific learning disabilities when an occupational therapist or psychologist is conducting a comprehensive battery of tests. Although there are limited screening tools for dysgraphia, they are not used as frequently as those for dyslexia. Screeners for dysgraphia, such as the Screener of Handwriting Proficiency, evaluate formation, size, neatness, speed, posture, pencil grip, and helping hand. However, these screeners only examine behaviors specific to dysgraphia and do not account for overlapping behaviors with dyslexia.

Challenges

Currently, no screeners examine the overlapping behavioral characteristics of dyslexia and dysgraphia in writing. To address this gap, the DDBIC was created. This screener examines behavioral indicators of dyslexia and dysgraphia in children's writing and provides specific actionable insights, particularly for a multi-tiered support system. It facilitates collaboration among educational practitioners to determine additional support based on the analysis of the child's writing sample. The DDBIC is being validated and normed on children in kindergarten through fifth grade. Another advantage of the DDBIC is its flexibility in analyzing a child's writing at any time and serving as a progress monitoring tool. Additionally, it can be used to analyze writing from daily academic tasks or more challenging writing prompts, such as narratives or expository texts [4].

Concerns

Although the DDBIC would be the first tool to use children's writing to screen for behavioral indicators of dyslexia and dysgraphia, scoring can be manually intensive, particularly when the ratio of practitioners to students is high. This problem is intensified by the fact that practitioners may have a limited understanding of how dyslexia and dysgraphia manifest in children's writing in the classroom [4]. The complexity of the indicators can also make identification difficult, further complicating the screening process. Practitioners may benefit from training to better understand how to identify behaviors and understand what behaviors are age-appropriate. Although this tool examines specific writing behaviors, the administration and scoring are done by hand. We are currently co-designing a system using a tablet and artificial intelligence to collect data and measure additional factors such as speed transition, pressure level, pen trajectory, and pen tilt, which have the potential to provide valuable insights into the screening process.

Another pressing concern is the lack of screening standards among current screening methods [5]. Without these standardized practices, the screening or evaluation process can vary significantly by region or school, lacking agreed-upon processes to ensure validity, accountability, or consistency. This could lead to erroneous recommendations from the screener. Additionally, there is also a lack of tools and guides adapted to the local language and culture.

Although tools are available, their availability and use differ considerably among more than 220 countries [5].

Summary

Early identification of behaviors related to characteristics of dyslexia and dysgraphia can significantly improve academic outcomes. For this to be successful, practitioners need an easy-to-implement screener that examines both conditions. This would provide insights into the underlying causes of the behaviors, allowing for appropriate instructional and interventional support. Furthermore, screeners should be quick and easy to administer and should be available for use whenever there is a concern. Artificial intelligence can allow scalable solutions that could revolutionize the screening process for dyslexia and dysgraphia, making it more efficient and accessible.

AI-Based Approaches to Screening

There is significant interest in leveraging artificial intelligence (AI) to boost the efficiency of the screening process. Thus, this section reviews the current AI-based systems for identifying the characteristics of dyslexia and dysgraphia through a child's handwriting. A summary of this comparative analysis is provided in Tables 4 and 5 for dyslexia and dysgraphia respectively.

Input Features

Common machine learning algorithms utilized for dyslexia and dysgraphia detection include support vector machines (SVMs), random forests (RFs), convolutional neural networks (CNNs), and AdaBoost [48]. These approaches utilize data that can be of two types: offline, which refers to scanned images of handwritten content on paper, and online, where a child writes on a digital tablet. Based on the type of data, different types of features are used for this detection task. Offline methods tend to utilize image features learned by

CNNs, as seen in [21, 32, 33, 47, 49]. Additionally, methods that automate existing screening tools, like [46], design algorithms to identify behavioral indicators specified by that tool.

On the other hand, online data consists of many features captured by the digital tablet. This tablet can obtain values such as the x-coordinate, y-coordinate, time, pen position, azimuth, altitude, and pressure [34]. These values can then be used to generate a variety of features that can be categorized into five types: dynamic, spatial, temporal, kinematic, and others. Dynamic features capture the pressure exerted while writing, altitude, and azimuth, while spatial features refer to the structural properties of one's handwriting, such as stroke length, width, and height. Temporal features consider properties like the time taken for writing and on-air time. Kinematic features explore the kinetic properties of writing, which include speed and acceleration. Other features refer to the number of pen elevations, erase count, and so on (Table 6).

Targeted SLD

Most AI systems that use handwriting samples tend to focus on dysgraphia detection rather than dyslexia detection due to the inherent characteristics and challenges associated with each SLD. Specifically, dysgraphia primarily affects the physical act of writing, including motor skills, letter formation, and spatial arrangement, which are directly observable and quantifiable from handwriting samples. These features can be effectively captured and analyzed using AI technologies to identify patterns indicative of dysgraphia. However, dyslexia is caused by difficulties with the phonological system and cannot be diagnosed by analyzing writing alone even though behavioral indicators can be seen in handwriting for children with dyslexia. To determine the cause and make a diagnosis of dyslexia, one would need to give an assessment that examines both oral and written language skills. Hence, a good number of systems that look at dyslexia focus on utilizing other

Table 4 List of child handwriting-based AI approaches for dyslexia detection

Reference	Nature of data	Dataset used	Features	Classifier	Classes
Spoon et al. [21]	Offline	[21]	CNN features	CNN	Dyslexic or not
Rosli et al. [22]	Offline	[23]	CNN Features	LeNet-5	Normal, reversal, or corrected handwriting
Alkhurayyif et al. [24]	Offline	[23]	YOLO V7 features	MobileNetV2-SSD lite	Normal, reversal, or corrected handwriting
Alqahtani et al. [25]	Offline	[23]	CNN Features	SVM, random forest	Normal, reversal, or corrected handwriting
Sreekumar et al. [26]	Offline	[23]	CNN features	CNN	Normal, reversal, or corrected handwriting
Jasira et al. [27]	Offline	[27]	CNN features	LSTM	Dyslexic or not
Isa et al. [28]	Offline	[28]	MLP features	MLP, OCR	Low risk or risk
Isa et al. [29]	Offline	[23]	CNN Features	CNN	Normal, reversal, or corrected handwriting

N/S, not specified; CNN, convolutional neural network; SVM, support vector machine, SSD, single shot detection, MLP, multilayer Perceptron

Table 5 List of child handwriting-based AI approaches for dysgraphia detection

Reference	Nature of data	Dataset used	Features	Classifier	Classes
Kariyawasam et al. [30]	Offline	[30]	Temporal, other	SVM	Dysgraphic or not
Devi et al. [31]	Offline	[31]	Handwriting, geometric features reclaimed from K-DCT-DTL	K-DCT-DTL	Dysgraphic or not
Hewapathirana et al. [32]	Online	[32]	CNN features	CNN	Dysgraphic or not
Kunhoth et al. [33]	Online	[33]	CNN Features	DenseNet201	Dysgraphic or not
Kunhoth et al. [34]	Online	[35]	175 features (dynamic, spatial, temporal, kinematic, and other)	AdaBoost	Dysgraphic or not
Mekyska et al. [36]	Online	[36]	51 features (spatiotemporal, kinematics, nonlinear dynamics, other)	Random forest	HPSQ score ^b
Asselborn et al. [37]	Online	[37]	53 features (static, tilt, pressure, kinematic)	random forest	Dysgraphic or not
Devillaine et al. [38]	Online	[38]	10 features (spatial, temporal, kinematics)	Random forest	Dysgraphic or not
Dankovičová et al. [39]	Online	[39]	52 features (spatial, temporal, kinematics, other)	SVM, random forest, AdaBoost	Dysgraphic or not
Deschamps et al. [40]	Online	[40]	40 features (static, tilt, pressure, kinematic)	SVM	Dysgraphic or not
Rosenblum et al. [41]	Online	[41]	21 features (spatial, temporal, kinematic, dynamic, other)	SVM	Dysgraphic or not
Sihwi et al. [42]	Online	[42]	Spatial, dynamic, temporal	SVM	Normal, light, moderate, or severe
Mekyska et al. [43]	Online	[43]	15 features (spatial, temporal, kinematic, dynamic, other)	XGBoost	With GD or not
Zvoncak et al. [44]	Online	[44]	TQWT applied on kinematic, and dynamic features	SVM, random forest	With HD or not
Zvoncak et al. [45]	Online	[45]	Spatial, temporal, kinematic	Gradient boosted tree	Dysgraphic or not
Drotár et al. [35]	Online	[35]	150 features (dynamic, spatial, temporal, kinematic, and other)	SVM, random forest, AdaBoost	Dysgraphic or not
Dimauro et al. [46]	Offline	[46]	13 BHK features	Algorithms	BHK score ^b
Ghouse et al. [47]	Offline	[35, 47]	R2CNN features	NDR-R2CNN	Dysgraphic or not

^aCNN, Convolutional Neural Network; SVM, Support Vector Machine; K-DCT-DTL, Kekre-Discrete Cosine Transform with Deep Transfer Learning; HPSQ, Handwriting Proficiency Screening Questionnaire; GD, Graphomotor Difficulties; TQWT, Tunable Q-factor Wavelet Transform; HD, Handwriting Difficulties; NDR-R2CNN, Non-Discrimination Regularization in Rotational Region Convolutional Neural network

^bThis is used to determine if the subject has dysgraphia or not by utilizing another classifier or a threshold

Table 6 List of child handwriting-based AI approaches for dyslexia-dysgraphia detection

Reference	Nature of data	Dataset used	Features	Classifier	Classes
Yogarajah et al. [49]	Offline	[49]	CNN features	CNN	Dyslexia-dysgraphia or not
Rashid et al. [50]	Offline	[50]	Spatial features	CNN	LD or NT

LD, learning difficulty (dyslexia and dysgraphia); NT, Neurotypical

modalities such as audio [30, 51–54] and facial [52, 55] features, eye-tracking [56–59], MRI [60–62], and EEG [63, 64] scans that can be obtained when a child is engaged in a reading task.

Identification Task

The focus of all of these approaches is on the detection (also referred to as diagnosis) of dyslexia and dysgraphia. Systems

of this nature are intended to perform binary classification, which involves determining whether or not the child has dyslexia or dysgraphia based on the provided handwritten sample. It is crucial to note that screening to determine whether a student has specific learning requirements differs from evaluating to diagnose a learning disability, as stated in [5]. Although a diagnosis may elucidate the underlying causes of a student's learning difficulties, it does not provide adequate information regarding the services that a student may require. This is because learning disabilities can have a diverse impact on individuals, even those who have received the same diagnosis. Rather, it is imperative to evaluate each student's unique learning strengths and challenges to identify the most suitable services and supports. Such information can aid teachers in creating an individualized education program (IEP) for the student, as well as determining what support the student will need to succeed in school.

Black Box Problem

Another limitation is that such ML-based diagnosis models are inherently opaque; it is challenging for humans to comprehend the model's reasoning behind the conclusion that resulted in its prediction. This lack of transparency, often referred to as the "black box" problem [21], means that practitioners are unable to understand how the model arrived at a particular diagnosis or decision. This information is essential for practitioners to provide ability-based interventions that are customized to the unique needs of each individual. Furthermore, without a clear understanding of the model's reasoning, it is difficult to fully trust its recommendations or adapt interventions effectively based on its outputs.

Summary

Thus, a review of the literature on AI systems for the identification of dyslexia and dysgraphia using handwriting samples has led to the following observations:

1. The majority of AI systems that utilize child handwriting samples are more focused on dysgraphia detection than dyslexia detection.
2. Most systems are intended to perform diagnosis rather than screening. Thus, they do not evaluate each student's unique learning strengths and challenges to identify the most suitable services and supports.
3. Current AI approaches for detection are inherently opaque, making it difficult for users to understand how the model arrived at a particular diagnosis. Such information is crucial for practitioners to provide ability-based interventions.

In light of this, we propose a framework describing different AI-based analytical modules to identify various behavioral indicators associated with dyslexia or dysgraphia from a child's handwriting sample in the "[Framework for Screening for Dyslexia and Dysgraphia](#)" section. The framework, which can be applied to any screening method, expedites the screening process, allowing practitioners to concentrate on a student's strengths and weaknesses to ascertain the necessary support and services, and the early administration of ability-based interventions.

Handwriting-Based Datasets for Screening

The availability and quality of datasets play a crucial role in the development and validation of AI-based approaches to the identification of dyslexia and dysgraphia. In this section, existing handwriting datasets are analyzed, along with a comparison between child and adult handwriting data. Table 7 lists some of the child handwriting datasets used for dyslexia and dysgraphia detection.

Nature of Data

As discussed in the "[AI-Based Approaches to Screening](#)" section, the nature of the data used is of two types based on the method of digitization: offline and online. Offline data consists of samples collected through the traditional pen-and-paper method which are then scanned to produce images. These types of data capture a range of static and spatial characteristics of handwriting, including the shape and structure of the written characters and the size of the strokes [33]. Online data, on the other hand, uses data written on digital tablets such as the Wacom Intuos digitizing tablet [33, 65], which is captured in the form of (x, y) coordinates and other features such as time, pen position indicator, azimuth, altitude, and pressure. Such data can be used to obtain dynamic, temporal, and kinematic features. In addition, online data can be converted into images to obtain static and spatial features similar to offline data.

Data Collection Task

Most dysgraphia-based datasets employ a copy task, even those that utilize writing activities based on current screening tools such as the BHK or ComPET. Although copy tasks are less strenuous on the cognitive load of a child, allowing them to focus more on the motor skills needed to write, they do not fully capture the cognitive and linguistic challenges associated with spontaneous writing [68]. Copy tasks primarily assess the physical act of writing and visual-motor integration, providing valuable insights into motor control and spatial arrangement. However, they fail to evaluate the child's ability to use appropriate vocabulary and grammar

Table 7 List of handwriting datasets used for dyslexia and dysgraphia detection

Reference	Grade/age	Availability	Nature of data	No. of adults	No. of children	Data collection task	Targeted SLD
Kariyawasam et al. [30]	Ages: 6–7	Private	Offline	0	100	Copy	Dysgraphia
Devi et al. [31]	Ages: 6–12	Private	Offline	0	290	N/S	Dysgraphia
Hewapathirana et al. [32]	Ages: 6–8	Private	Offline	0	N/S	Copy	Dysgraphia
Kunthoth et al. [33]	Ages: 8–15	Public	Online	0	120	Copy	Dysgraphia
Drotár et al. [35]	Ages: 8–15	Public	Online	0	120	Copy	Dysgraphia
Mekyska et al. [36]	Ages: 8–9	Private	Offline	0	54	CompPET	Dysgraphia
Asselborn et al. [37]	Grades: 1–5	Private	Online	0	298	BHK	Dysgraphia
Devilleine et al. [38]	Grades: 2–5	Private	Online	0	300	Drawing	Dysgraphia
Dankovičová et al. [39]	Ages: 10–13	Private	Online	0	78	Copy	Dysgraphia
Deschamps et al. [40]	Grades: 2–5	Private	Online	0	580	BHK	Dysgraphia
Rosenblum et al. [41]	Ages: 8–9	Private	Online	0	99	Copy	Dysgraphia
Sihwi et al. [42]	Grades: 3–6	Private	Online	0	32	Write random sentences	Dysgraphia
Mekyska et al. [43]	Grades: 1–4	Private	Online	0	76	Copy of specific drawings	Dysgraphia
Zvoncak et al. [44]	Grades: 3–4	Private	Online	0	65	Copy	Dysgraphia
Zvoncak et al. [45]	Grades: 3–4	Private	Online	0	97	Copy	Dysgraphia
Dimauro et al. [46]	Grades: 2–5	Private	Offline	0	109	BHK	Dysgraphia
Ghouse et al. [47]	Ages: 6–8	Public	Offline	0	150	N/S	Dysgraphia
Gemelli et al. [66]	Ages: 9–10	Private	Online	103	95	BHK	Dysgraphia
Asselborn et al. [67]	Ages: 5–12	Available on request	Online	0	448	BHK	Dysgraphia
Spoon et al. [21]	Grades: K–6	Private	Offline	0	88	Existing work, Copy, Dictation, Composition with prompt	Dyslexia
Isa et al. [23] ^b	N/S	Public	Offline	N/S	N/S	Collection of children's samples and other public datasets	Dyslexia
Isa et al. [28]	Ages: 7–12	Private	Offline	0	30	N/S	Dyslexia
Jasira et al. [27] ^c	N/S	Private	Offline	0	N/S	N/S	Dyslexia
Yogarajah et al. [49]	Grades: 1–5	Private	Offline	0	54	Existing work	Dyslexia, Dysgraphia
Rashid et al. [50]	Ages: 5–15	Private	Offline	0	25	Copy, Write days and months, Composition with pictorial cue	Dyslexia, Dysgraphia

^aBHK, brave handwriting kinder; HPSQ, handwriting proficiency screening questionnaire; CompPET, computerized penmanship evaluation tool; N/S, not specified

^bNumber of samples in the dataset is 138,500

^cNumber of samples in the dataset is 86,115

and construct coherent sentences. These aspects are also crucial for comprehensively understanding dysgraphia, as the condition can impact higher-order writing skills as well. Therefore, while copy tasks are useful for isolating and examining motor skill deficits, a more holistic approach that includes tasks requiring spontaneous writing and composition is necessary to assess and fully address the multifaceted nature of dysgraphia [69–71].

In addition to copying tasks, the dyslexia-based dataset designed by Spoon et al. [21] incorporates other activities such as dictation and composition in addition to collecting a child's existing written work. However, Rosli et al. [22] propose a dataset consisting of just letters, which presents several significant limitations as it does not capture the full spectrum of difficulties experienced by individuals with dyslexia. While analyzing letters can provide insights into visual-motor integration and basic phonological processing, it overlooks higher-order language skills such as sentence construction, grammar, and coherent thought organization. This narrow focus can miss the nuances of dyslexia that manifest in complex reading and writing tasks. Additionally, letters alone do not reflect a child's ability to comprehend and generate meaningful text, which is crucial for a comprehensive assessment of dyslexia. This approach also fails to account for contextual usage and the impact of dyslexia on spelling and vocabulary.

Comparison with Adult Handwriting

Numerous adult handwriting datasets, such as IAM [72], RIMES [73], and CVL [74], are readily accessible online. However, they are not suited for developing dysgraphia and dyslexia screening models due to the absence of the distinctive characteristics found in children's handwriting. Adult samples typically lack structural irregularities like incomplete letter formation, crowded writing, and improper letter placement, which are crucial indicators of dysgraphia and dyslexia in children. Therefore, for effective dysgraphia and dyslexia screening, it is essential to use models trained specifically on child handwriting samples that exhibit these features. Unfortunately, as previously mentioned, there is a significant shortage of comprehensive and structured child handwriting datasets that are publicly available, which hampers the development of effective screening tools. The lack of child datasets impedes the development of robust screening tools, underscoring the need for targeted datasets that capture the unique aspects of child handwriting.

Summary

In essence, the following observations can be made from our review of child handwriting datasets:

1. Existing child handwriting datasets catered towards dyslexia are mostly offline. In contrast, dysgraphia-based datasets are mainly online.
2. Most child handwriting datasets are not publicly available.
3. There are only a few handwriting datasets geared toward dyslexia analysis.
4. The majority of handwriting-based models are trained and evaluated on adult handwriting datasets, which are readily accessible. These models may not be compatible with child handwriting, as discussed in the “[Framework for Screening for Dyslexia and Dysgraphia](#)” section.

To mitigate these issues, we propose to build a dataset specifically for dyslexia and dysgraphia that consists of both offline and online samples written by children. Such data will be collected through the methodology presented in the “[Methodology for Data Collection](#)” section with the aid of practitioners such as SLPs and general educators. This data can be used to build AI-based tools that can identify behavioral indicators present in the child's handwriting and screen for dyslexia and dysgraphia.

Scoping Review Strategy

A scoping review was conducted to explore the existing literature relevant to the screening of dyslexia and dysgraphia, particularly focusing on the integration of artificial intelligence (AI). This method allowed us to identify key concepts and highlight gaps in the literature.

The research questions guiding this review were as follows:

1. What are the current methods practitioners employ to screen children for dyslexia and dysgraphia, and what challenges do they encounter?
2. How can AI be utilized to expedite the screening process for these disabilities?
3. What are the existing AI approaches for screening dyslexia and dysgraphia?
4. How effective are these AI systems in screening? Do they complement current screening practices?
5. What is the extent of practitioner involvement in these AI systems?
6. Are there publicly available child handwriting datasets in the literature? If so, what are their characteristics?

To address these questions, a comprehensive search was conducted across multiple databases, including Google Scholar, PubMed, ERIC, Academic Search Premier, Web of Science, APA PsycInfo, and CINAHL Complete. Keywords and search terms were selected to encompass a wide range of studies, including “handwriting”, “dyslexia”,

“dysgraphia”, “detection”, “diagnosis”, “screening tools”, “AI”, “deep learning”, “machine learning”, “children”, “letter formation”, “letter crowding”, “spacing”, “legibility”, “quality”, and “spelling”. The titles and abstracts of papers encountered during the preliminary search were screened for relevance to the research questions. The findings provide a foundation for understanding the domain and identifying critical gaps that can be addressed with our proposed framework.

Methodology for Data Collection

This section discusses our proposed approach to data collection to mitigate the challenges discussed in Sect. “Handwriting-Based Datasets for Screening”. This method is co-designed with speech-language pathologists (SLPs), occupational therapists (OTs), special educators,

and general educators. A study protocol incorporating this strategy has been approved by the Institutional Review Board (IRB) at the University of Nevada, Reno for the collection of data at an elementary school near Reno. Figure 2 illustrates this approach.

Given the scarcity of useful data sets with children’s writing, and the varying challenges posed by offline and online data, as well as the complementary information that online data can provide, the proposed effort focuses on the collection of both offline and online handwriting samples from children spanning the elementary school age range. As discussed earlier, offline samples are those collected through traditional pen-and-paper methods, whereas online samples are obtained using a digital tablet. With our data collection approach, we hope to develop datasets that can provide better insights into understanding the writer’s cognitive processes and motor skills during the writing process.

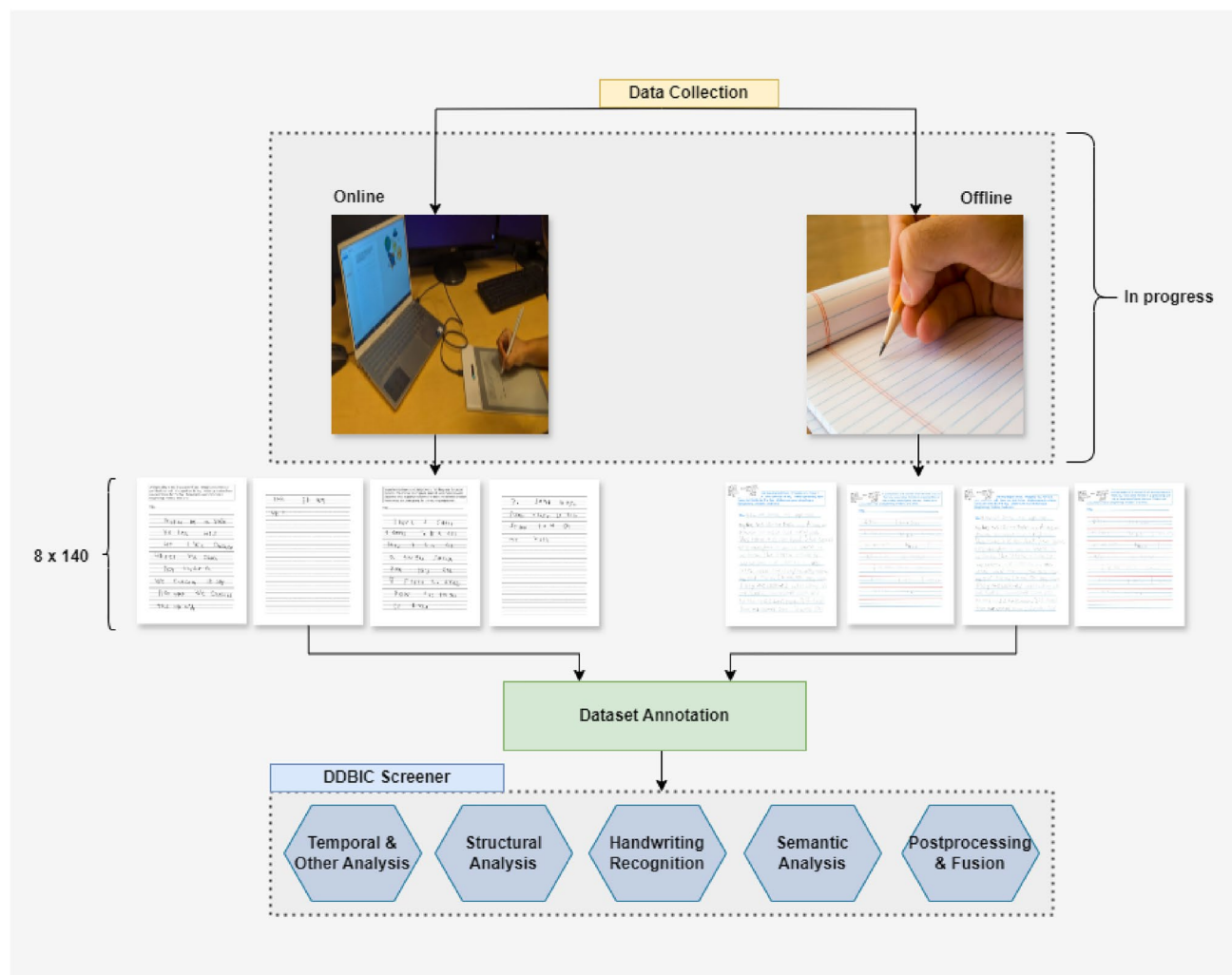


Fig. 2 An overview of the data collection process

Collection of Handwriting Samples

Our ongoing study aims to collect students' writing samples to inform the psychometric properties (validity and reliability) of the Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC), while also creating data sets for developing and testing algorithms for automatically extracting multiple modality metrics of behaviors and writing and benchmarking of current state-of-the-art algorithms and improving their performance using this dataset. We also intend to assess whether AI-coded measures of nonverbal communication behaviors and writing are comparable to human-coded measures.

Students in grades K through 5th grade are included in this research, as the DDBIC is designed to identify behavioral characteristics of dyslexia and dysgraphia in kindergarten through 5th-grade school-age children. The participant population includes students with or without an individualized education program (IEP) who can complete a writing task in English and be of any gender or ethnic background. The projected sample size is expected to be 130–260 students, where a power analysis indicates that a minimum of 20 students per grade would be needed to statistically analyze each of the 13 behavioral indicators on the DDBIC.

Parents/guardians will be invited to have their children participate in the study. In addition, permission to video record the session will also be sought from them. Consent for students to participate in the study and student assent for children 7 years and older will be collected at the beginning of each data collection session as approved by the IRB. The entire procedure will be thoroughly explained to the child, ensuring they understand that participation is voluntary and they are free to withdraw at any time without any consequences. Additionally, each child will be assigned a distinctive numerical identifier to ensure that the collected samples are de-identified to preserve privacy.

Each child will participate in data collection across two time points, which are within 1 day to 3 weeks apart. During a typical data collection session, each child will participate in four activities: two using paper and two using a tablet. Each activity will entail writing a response to a story starter prompt with a visual cartoon-like picture cue. There will be six story themes, each having a narrative and expository prompt. These prompts will guide a child to write a narrative or expository response, which is intended to elicit a variety of writing styles. Story prompts will be counterbalanced across genres and story prompt themes to ensure an equal amount of stories across grades. This method will guarantee a diverse and exhaustive collection of handwriting samples, which will capture the subtleties of each child's writing abilities in a variety of mediums and contexts.

Need for Narrative and Expository-Based Prompts

Narrative and expository prompts are selected to mirror academic activities for school-aged children. Prompts that elicit narrative or expository responses are valuable tools in the analysis of a child's cognitive thinking and development. These types of prompts will stimulate different cognitive processes and thinking capabilities of a child when used in handwriting analysis. Narrative-based prompts will ask children to create a story or a personal experience and evaluate a child's ability to sequence events, their capacity for creative thinking, and their ability to link ideas logically. Conversely, expository writing requires the child to explain a concept or argue a point of view. These prompts assess skills like the ability to organize information logically and the use of language to inform or persuade.

The following factors are also important in driving the use of narrative and expository-based prompts in assessments:

1. Most of the existing research utilizes datasets comprised of letters or words written by children [21, 23, 28, 30, 32, 33, 39, 49, 50]. However, these data often fail to capture structural characteristics, such as word spacing or alignment with respect to the left margin. Moreover, such datasets employ activities where children are merely copying words, sentences, or paragraphs [21, 30, 32, 35, 39, 41, 43–46]. These tasks neither provide adequate complexity or challenge to children nor do they sufficiently engage a child's cognitive processes.
2. Dysgraphia and dyslexia primarily impact academic skills and manifest prominently in educational settings. Hence, it is crucial to conduct screenings in environments that replicate the academic conditions a child encounters daily in school. This setup helps in accurately assessing a child's cognitive and analytical skills in a realistic context.
3. Each U.S. state has established academic standards for skills in multiple disciplines that students are expected to meet at certain grades. This includes state standards for narrative and expository stories [75]. Thus, employing narrative and expository prompts allows for a direct evaluation of how well students meet these specific educational criteria.

Thus, the use of narrative and expository-based prompts will ensure a diverse and comprehensive set of handwriting samples, capturing the nuances of each child's writing abilities across different mediums and contexts.

Post-collection

Statistical Analysis of Data

To gain a deeper understanding of the behavioral indicators associated with dyslexia and dysgraphia, as well as their overlap, the data collected will undergo a multi-step analysis process. This analysis will focus on the following areas: (i) manual verification of the behavioral indicators; (ii) validation of results using established screening tools such as TILLS SLS; (iii) analysis of the occurrence of behaviors across different grades and modalities; and (iv) the identification and validation of online behaviors that cannot be measured on paper. SLPs and OTs will manually verify behavioral indicators and concurrently validate the DDBIC using existing methodologies such as the TILLS SLS. OT evaluations will be conducted to identify children who are at minimal risk, moderate risk, or significant risk of one or both of these SLDs. The analysis will also be stratified by grade level, with means and standard deviations calculated for each grade and comparisons made across modalities, that is, paper versus tablet. Additional checks to identify and validate online behaviors, such as pauses and writing pressure that cannot be measured manually, will be performed using an AI-integrated screener, as detailed in the “[Framework for Screening for Dyslexia and Dysgraphia](#)” section.

Data Annotation

Data annotation entails labeling each sample with relevant metadata, such as handwriting style, text transcriptions, and any observed errors, to ensure the dataset is comprehensive and usable for various analytical purposes. Pre-processing steps, such as digitizing handwritten texts and normalizing the data to a consistent format, are essential to preparing the dataset for analysis. Next, detailed annotation is performed, which may include labeling individual characters, words, and sentences, as well as marking features such as legibility, alignment, and spacing. Expert observations by SLPs on the quality of idea development and organization in the writing samples will also be recorded. Additionally, expert reviewers might be required to ensure the accuracy and consistency of annotations. Quality control measures, such as inter-annotator agreement checks, help maintain the reliability of the dataset. Finally, the annotated dataset should be organized and stored in a structured format with clear documentation to facilitate its use in the research and development of the component technology modules and the screening tool.

Framework for Screening for Dyslexia and Dysgraphia

This section describes the proposed framework for developing an AI-integrated screening tool designed to enhance the manual evaluation process by identifying various behavioral indicators from children’s handwriting samples. The tool aims to assist practitioners by providing explainable and actionable insights, streamlining the traditionally time-intensive and subjective process of handwriting analysis. We discuss this framework in the context of the DDBIC screener which is a compendium of behavioral characteristics identified by practitioners in extant speech-language pathology literature. The Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC) [4] is utilized to screen for dysgraphia and dyslexia by analyzing child handwriting. The framework integrates model-driven principles and a data-driven approach in a complementary manner. Model-driven principles, as outlined in the DDBIC, provide a foundation by defining key behavioral indicators such as spatial misalignment, inconsistent letter formations, and limited vocabulary. These indicators are based on expert knowledge from speech-language pathology and occupational therapy, ensuring the framework is aligned with established practices used by practitioners.

Building on this foundation, the framework employs a data-driven approach to enhance the scalability and efficiency of handwriting analysis. By training machine learning models on diverse handwriting datasets, the system learns patterns and features that are indicative of these learning disabilities. The integration of model-driven and data-driven approaches enables the framework to effectively analyze handwriting data at scale, leveraging machine learning’s capabilities while ensuring the results are rooted in expert-defined criteria. This synergy bridges the gap between theoretical knowledge and practical application, creating a system that is both scalable and interpretable.

The DDBIC [4] is comprised of 17 behavioral indicators that can be observed both during and after the writing process, as illustrated in Fig. 3. Practitioners can observe four key indicators while a child writes, including pencil grip, handwriting pressure, writing speed, and behaviors such as self-esteem, task engagement, and avoidance behavior. These indicators are monitored over time to assess their influence on the child’s writing and developmental progress. Practitioners may evaluate an additional 13 behavioral indicators after the child has finished writing, and these are divided into three categories: (a) dysgraphia exclusively, (b) dysgraphia and dyslexia, and (c) dyslexia exclusively. The DDBIC specifies five distinct indicators of dysgraphia: inconsistent letter scaling, inversions, transpositions, atypical use of margins or space planning, and atypical

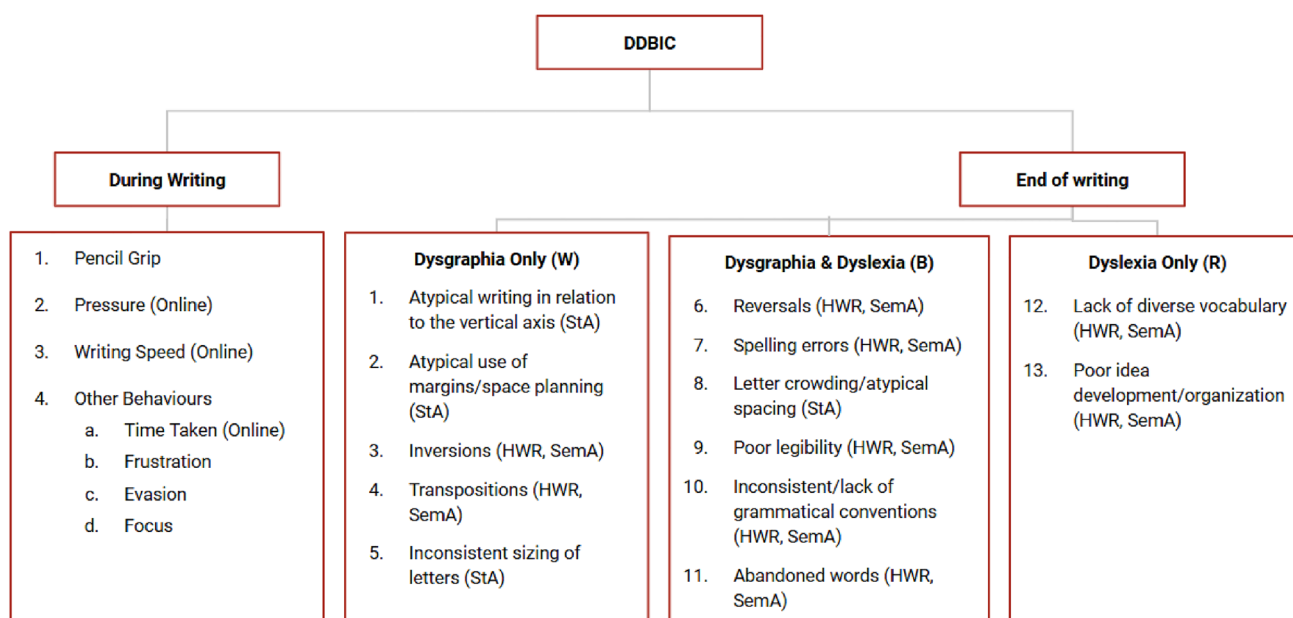


Fig. 3 Overview of the behavioral indicators as outlined in the Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC)

writing relative to the vertical axis. Six indicators may indicate dysgraphia and/or dyslexia, including reversals, spelling errors, letter congestion or unusual spacing, poor legibility, inconsistent or absent grammatical conventions, and abandoned words. Furthermore, dyslexia is explicitly linked to two indicators: inadequate idea development or organization, and a restricted vocabulary.

The identification of DDBIC behavioral indicators in a child's handwriting will require the integration of multiple functional modules. These modules are tailored to address different aspects of handwriting analysis and include: (1) Temporal and Other Analysis (TA), (2) Structural analysis (StA), (3) Handwriting Recognition (HWR), (4) Semantic Analysis (SemA), and (5) Post-processing and Fusion. Each module captures and analyzes distinct characteristics of handwriting, making a unique contribution to the overall assessment. Figure 3 maps the analytical modules that will be required to capture each specific DDBIC behavioral indicator.

Temporal and Other Analysis (TA)

The Temporal and Other Analysis (TA) module examines behavioral indicators that possess temporal attributes and dynamic elements that are specific to the writing process. These include *writing speed*, *sequence of character formation*, *presence of pauses*, and *pen tilt*, all of which can be observed during writing. Supplementary factors such as *writing pressure*, *pen trajectory*, and *pen grip*, may

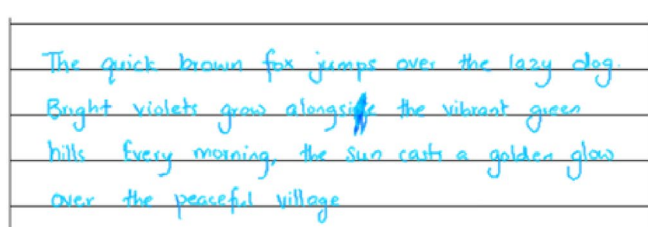
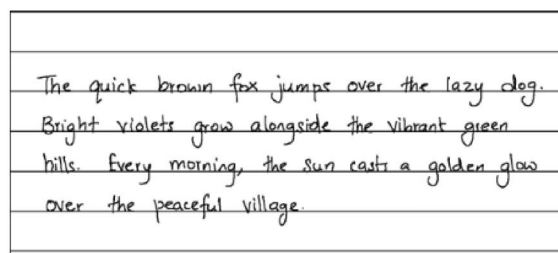
also provide valuable insights into the cognitive and motor processes that underlie handwriting.

Pressure reflects the force exerted between a child's hand and the writing instrument during writing. This metric is easier to measure in an online sample collected using a tablet, which can directly capture quantitative pressure data. Here, pressure sensitivity is gauged by detecting the various levels of force applied to the stylus. In contrast, analyzing pressure in offline data is more challenging but can be inferred by examining the intensity or color variations in the ink of the written letters or by identifying physical impressions such as tearing of the paper, which may suggest the application of excessive pressure while writing. Measuring pressure levels can provide insights into the stress level and potential difficulties a student may be facing while writing. Higher and lower pressure levels might indicate increased stress or discomfort, suggesting challenges with writing [4].

In the online data collection modality, additional handwriting-based dynamic measurements such as writing speed, pen tilt, pen trajectory can be quantifiably obtained from the hardware. *Writing speed* is determined by tracking the time it takes for the stylus to move from one point to another, *pen tilt* is measured by the angle at which the stylus is held, and *pen trajectory* tracks the precise path of the stylus. Online sample collection facilitates the capture of the unique patterns of an individual's handwriting and drawing style. Figures 4 and 5 illustrate a sample collected using the Wacom IoT Paper tablet and the metrics obtained from the handwritten sample, respectively.

Fig. 4 A writing sample captured on the Wacom tablet. To the left is a summary of the online features obtained

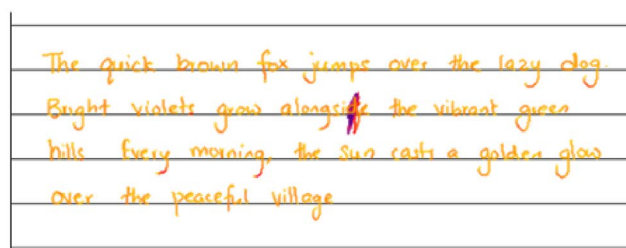
Writing time: 103 sec
 Stroke count: 168
 Pen down ratio: 46%
 - Pen speed ave.: 16 mm/s
 (Max: 151 mm/s Min: 0 mm/s)
 - Pressure ave.: 704
 (Max: 4095 Min: 155)
 - Packet count: 40193
 (On: 22965 Off: 17228)
 - Stroke interval ave.: 332 msec
 (Max: 2443 Min: 33)



Speed 0 to 300 mm/s



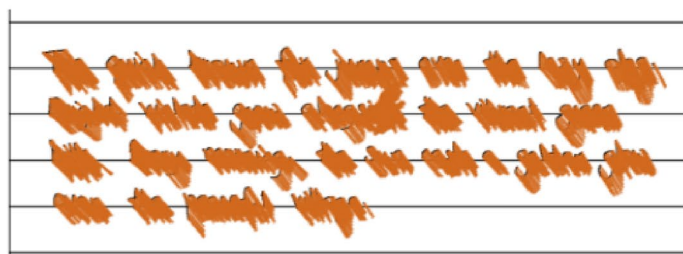
(a)



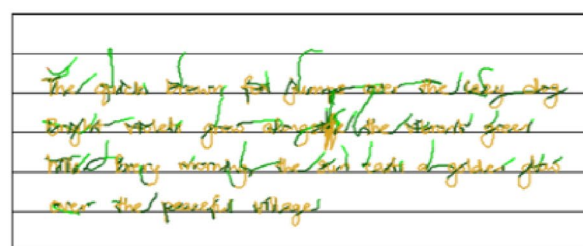
Pressure 0 to 4096



(b)



(c)



Height 0 to 10 mm



(d)

Fig. 5 Online features obtained from the writing sample in Fig. 4: **a** writing speed **b** pressure **c** pen tilt **d** pen trajectory

We anticipate that the rich and varied data being captured in the proposed dataset will allow us to analyze the student's temporal handwriting dynamics, and reliably identify behaviors that can contribute to a robust screening process for dyslexia and dysgraphia.

Structural Analysis (StA)

The Structural Analysis (StA) module evaluates spatial indicators such as *letter crowding/atypical spacing*, *inconsistent sizing of letters*, *atypical use of margins*, and

atypical writing in relation to vertical axis. These indicators can be evaluated in both offline and online samples after the completion of a writing task. Typically, margins should be vertically aligned, letters should be uniformly sized, spacing between letters should be consistent, and letters should align with the horizontal line. Deviations from these typical conditions are assessed through structural analysis, providing insights into the specific behavioral indicators present in the handwriting sample.

This module will leverage text detection, line detection, and/or character detection algorithms to accurately extract structure and layout information from images of handwriting samples. A combination of these structural building blocks can be leveraged to develop a reliable *atypical use of margins* indicator, in consultation with practitioners. Variations in the use of margins may occur in typical children's handwriting as well and a deeper understanding of how practitioners evaluate atypical use of margins that may be indicative of a potential problem will drive the design of this module. Figure 6 provides examples of atypical handwriting. In both samples shown in the figure, the handwriting demonstrates significant deviations from typical patterns. The text is misaligned with the horizontal writing line, with letters placed both above and below it, and the spacing between letters is inconsistent. In Fig. 6a, the child intended to write "I want to jump on the moon." However, the spacing between the words "on" and "the" in the third line is inconsistent, resulting in insufficient word separation. This causes text detection algorithms to misinterpret these as a single word. While the vertical alignment with respect to the left margin is relatively good, the alignment with the right margin is poor. For example, the words "on" and "the" on the third line could have fit within the second line, indicating inefficient use of space. In Fig. 6b, the writing displays multiple atypical characteristics. The spacing between letters and words is inconsistent, resulting in areas of crowded writing.

Most letters are positioned below the horizontal writing line, and the spacing from the left margin is uneven. To detect and analyze these behavioral indicators in the AI-enhanced screener, advanced structural analysis techniques must be employed. These include text and line detection algorithms capable of identifying deviations in spacing, alignment, and letter positioning.

Line segmentation and detection algorithms [76–78], as well as text detection algorithms like Character Region Awareness for Text Detection (CRAFT) [79] and Efficient and Accurate Scene Text Detector (EAST) [80], can effectively find text or line areas in handwriting samples and draw boundaries around them. Additionally, algorithms that utilize horizontal and vertical histograms on the pixel matrix are effective in detecting lines and words in handwritten images [46]. The distance from the left and right margins can be determined by the coordinates of the first and last bounding boxes in each line, which are detected using a text detection algorithm. These distances can also be calculated based on the left and right coordinates of the bounding box that encloses each line, which are detected using a line detection algorithm. Figure 7a illustrates how the CRAFT algorithm detects text and calculates the distance from the left and right margins (denoted as LD and RD) relative to the position of the bounding boxes, which are close to the margins at each line. Here, the distances from the left and right margins vary across different lines, indicating *atypical use of margins*. The slope of the bounding boxes can help characterize the skew of texts in the writing sample.

The evaluation of indicators such as *letter crowding*, *atypical spacing*, *inconsistent sizing of letters*, and *atypical writing in relation to vertical axis* requires detailed character-level text detection and analysis. A variety of techniques are available in the literature for character detection/segmentation in images of handwriting samples [81–83]. Most character detection algorithms yield suboptimal results

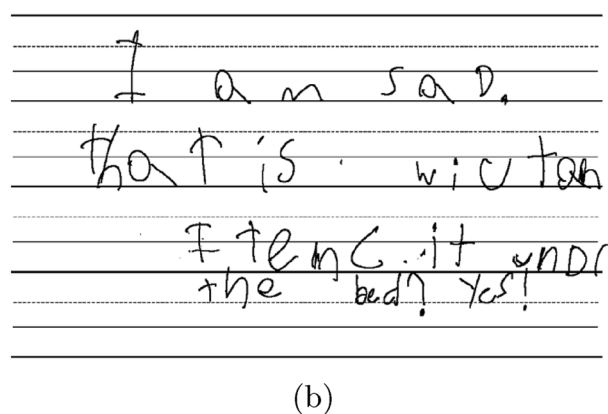
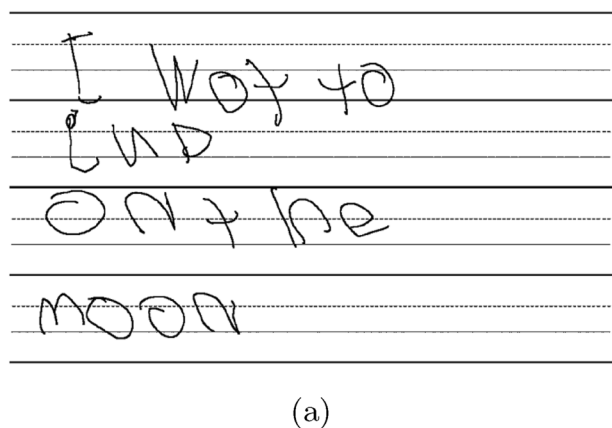


Fig. 6 Handwriting samples with atypical writing **a** from kindergarten, **b** from second grade

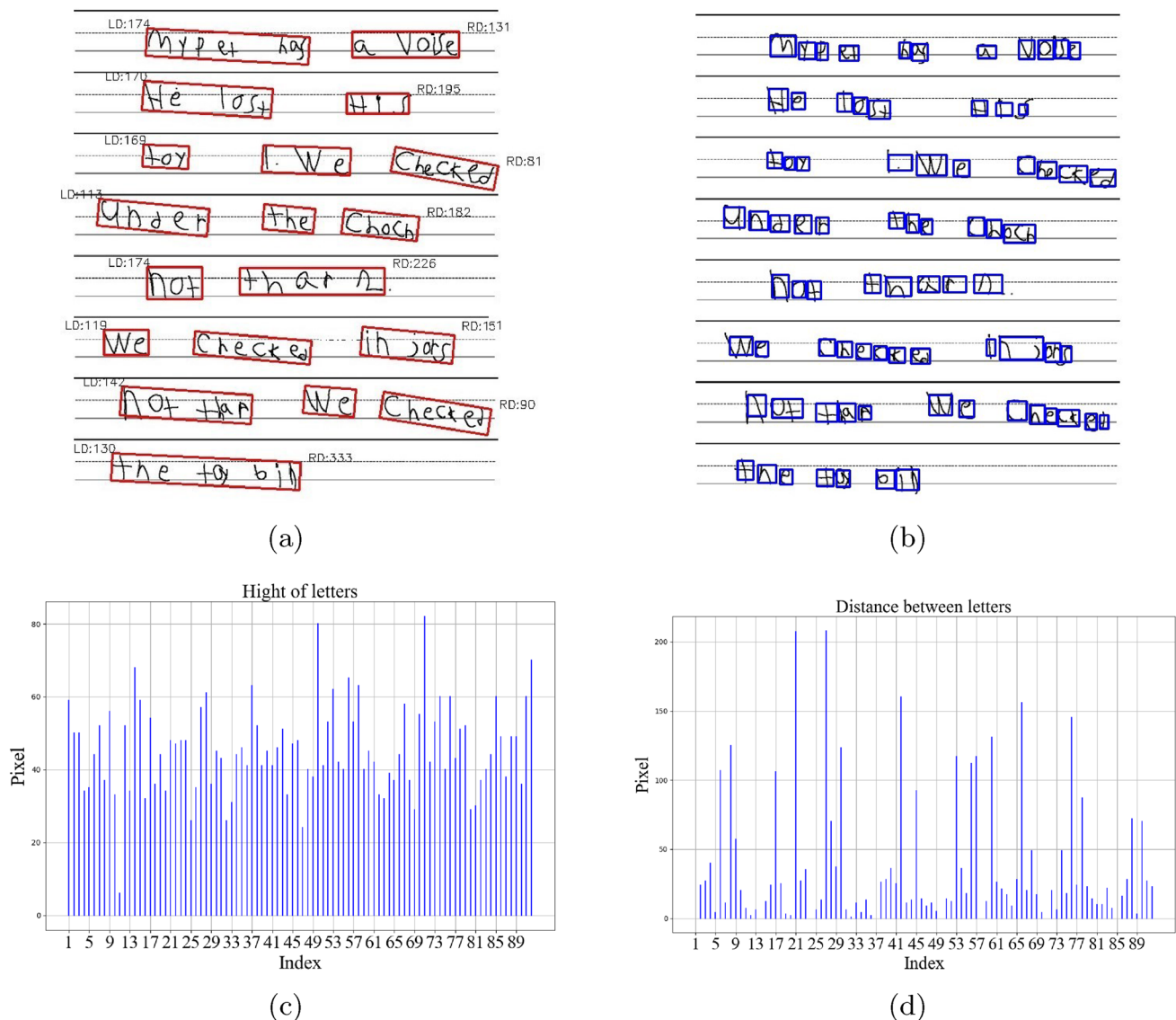


Fig. 7 Structural analysis of handwriting samples **a** atypical margin use, **b** inconsistent letter sizing **c** histogram representing height of characters **d** histogram representing distance between characters

on child handwriting due to irregularities and varying spacing between characters. Figure 7b describes the result of a character segmentation algorithm. We use the heat maps generated by CRAFT [79] for character segmentation, followed by the application of a contour detection algorithm to enclose letters within bounding boxes.

The height of the generated bounding boxes reflects the size of each character. Bounding boxes of varying heights indicate *inconsistent sizing of letters*. Furthermore, the space between consecutive bounding boxes can reveal aspects of the writing's spacing: wide gaps suggest sparse writing, whereas overlapping boxes point to congested writing. The distance between the coordinates of two adjacent bounding boxes shows the spacing between characters. Uneven

spacing between the character bounding boxes within a word indicates *atypical spacing* in the writing. The distance from the lower edges of the character bounding boxes to the writing line is crucial for evaluating the alignment of letters relative to the baseline. Inconsistent positioning of characters above or below this horizontal writing line indicates *atypical writing in relation to the vertical axis*. This particular aspect of handwriting can be analyzed by comparing the lower coordinates of the character bounding boxes with those of the horizontal writing line, providing a quantitative assessment of vertical alignment discrepancies. Figure 7c presents a histogram illustrating the height of letters in the child's handwriting samples, where the variation in line heights reflects inconsistent letter sizing.

Similarly, Fig. 7d displays a histogram of the distances between letters, with irregularities in the lines indicating atypical spacing.

In summary, the StA module plays a vital role in evaluating spatial indicators by building on a combination of line/word/character segmentation algorithms to analyze both offline and online handwriting samples. The challenge will be to identify structural deviations from typical writing for a child at a particular grade level and reliably identify handwriting characteristics that could suggest underlying learning challenges.

Handwriting Recognition (HWR)

The Handwriting Recognition (HWR) module is crucial for the textual analysis of handwriting samples, as its primary goal is to transcribe what the child has written accurately. However, identifying certain behavioral indicators, such as spelling errors, letter reversals, and letter inversions, often requires a collaborative effort between the HWR module and the Semantic Analysis (SemA) module. This synergy is essential for understanding not only the actual written content but also the child's intended meaning. For instance, in Fig. 6a, the child wrote (preserving the original letter cases and spaces): "I wot to juD onthe moon," where the letter "j" is reversed, and the words "wot" and "juD" are misspelled. The intended text, however, is "I want to jump on the moon." While the spellings are incorrect, they represent plausible phoneme-to-grapheme conversions (e.g., "wot" for "want" and "juD" for "jump"). Such errors might not be considered significant indicators of dyslexia or dysgraphia in lower grades due to developmental spelling patterns. However, in higher grades, these same errors could become more critical indicators of learning disabilities. An efficient OCR system is needed to accurately transcribe both the actual and intended text in these scenarios, capturing nuanced errors like letter reversals and unconventional spelling. In contrast, Fig. 6b presents errors that deviate from typical phoneme-to-grapheme conversion. The child wrote: "I am saD that is wiutan ItemC it unDr the bed? Yes!" (preserving letter cases and spaces). However, their intended transcription was: "I am sad that is why we are going to find it. Is it under the bed? Yes!" This discrepancy between the actual and intended writing highlights a more serious indicator, as the intended meaning and structure significantly diverge from the written text. Such deviations are not developmental spelling errors but may reflect deeper cognitive or linguistic challenges.

A variety of optical character recognition (OCR) technologies are available, including commercial services such as Microsoft Azure OCR [84], Amazon Textract [85], and open-source pipelines such as PaddleOCR [86], EasyOCR [87], Tesseract [88], and TrOCR [89]. However,

these OCR technologies are predominantly trained on printed text or adult handwriting, which often results in difficulties in accurately processing children's handwriting, especially that of younger children due to their widely differing style, generally characterized by poor legibility and lack of structure.

Figure 8 illustrates the result of Microsoft Azure OCR on two handwriting samples: a recently collected sample from our child data collection and an adult sample obtained from the IAM dataset [72]. Specifically, Fig. 8c illustrates the OCR results for the child's handwriting sample shown in Fig. 8a, while Fig. 8d illustrates the OCR results for the adult's handwriting sample shown in Fig. 8b. The recognition accuracy for the child's handwriting is notably lower than the adult's, with the character error rate (CER) for the adult sample at 1.75% and the child sample at 3.48%. This discrepancy is likely due to the poor developmental nature of a child's handwriting, and features such as poor legibility, incomplete letter formation, crowded writing, improper placement of letters above or below the writing line, letter reversals, and transpositions found in a child's handwriting can contribute towards inadequate recognition outcomes with current OCR technologies. This underscores the need for more sophisticated OCR solutions that are specifically designed to accommodate the intricacies of children's handwriting. Figure 9 illustrates the recognition results of child handwriting using PaddleOCR, EasyOCR, and a hybrid method that combines CRAFT with TrOCR, with CERs of 16.27%, 33.33%, and 10.85%, respectively, underscoring the need for a more adaptable recognition system. These results highlight the necessity for further fine-tuning of existing OCR technologies on children's data and the development of additional metrics to better understand children's handwriting and enhance the accuracy and reliability of these systems.

Furthermore, handwriting recognition in offline samples presents a variety of additional challenges in comparison to samples collected using online methods. One significant issue is the poor visual quality and variability in handwriting, which can considerably impede the automated recognition of handwriting. Figure 10 illustrates the OCR results on an offline-collected handwriting sample, where the system struggles to detect some line segments due to faint text. Thus, it may necessary to employ preprocessing techniques to enhance the clarity and quality of these offline samples before further analysis can be conducted. Techniques such as image enhancement and binarization, morphological operations, and noise removal can help improve visual quality, leading to better prediction.

Some of the behavioral characteristics exhibited by children with dyslexia and dysgraphia include reversals and inversions as well as characters that may not be well formed, which results in symbols that are not classes in typical

Animals play a role in people's lives. Imagine you have a
pet that can talk. Your pet loses its toy. Make up a story, have
your pet look for the toy. Make sure your story has a
beginning, middle, and end.

Title _____

My pet has a voice
He lost HIS
toy 1. We checked
under the Choch
not that 2.
We checked in jors
We checker
the fo bil

(a)

My pet has a voice
He lost HIS
toy 1. We checked
under the Choch
not that 2.
We checked in jors
We checker
the fo bil

(c)

Sentenced Database

401-28

When Mr. Brown sat down Labour MP. He cheered for a full minute - and even his bitterest opponents on defence joined in. Mr. Powell devoted half his speech to giving details of plans for improving the hospital service, on which indeed the Government is making progress. His basic defence of the Health Service cuts was that "even after the proposed changes the net cost of the service to the Exchequer will have increased over three years by 20 per cent."

When Mr. Brown sat down Labour MP. He cheered for a full minute - and even his bitterest opponents on defence joined in. Mr. Powell devoted half his speech to giving details of plans for improving the hospital service, on which indeed the Government is making progress. His basic defence of the Health Service cuts was that "even after the proposed changes the net cost of the service to the Exchequer will have increased over three years by 20 per cent."

Name _____

(b)

When Mr. Brown sat down Labour MP. He cheered for a full minute - and even his bitterest opponents on defence joined in. Mr. Powell devoted half his speech to giving details of plans for improving the hospital service, on which indeed the Government is making progress. His basic defence of the Health Service cuts was that "even after the proposed changes the net cost of the service to the Exchequer will have increased over three years by 20 per cent."

(d)

Fig. 8 Recognition of online handwriting samples using Microsoft Azure OCR, **a** child sample, **b** adult sample, **c** recognition result of child sample (a), **d** recognition result of adult sample (b)

OCR systems. *Inversion* refers to letters flipped along the horizontal axis, and *reversals* refer to the flipping of letters along the vertical axis. Recognizing these characteristics will require retraining the OCR with additional classes representing these unique symbols.

Thus, the HWR module for our framework will require adapting state of the art OCR technologies to address the

significant challenges of children's handwriting. The effectiveness of current OCR technologies is reduced by the poor legibility and atypical writing styles of children, underscoring the need for OCR solutions specifically designed for child handwriting. Specialized pre-processing techniques and OCR algorithms will be necessary to improve

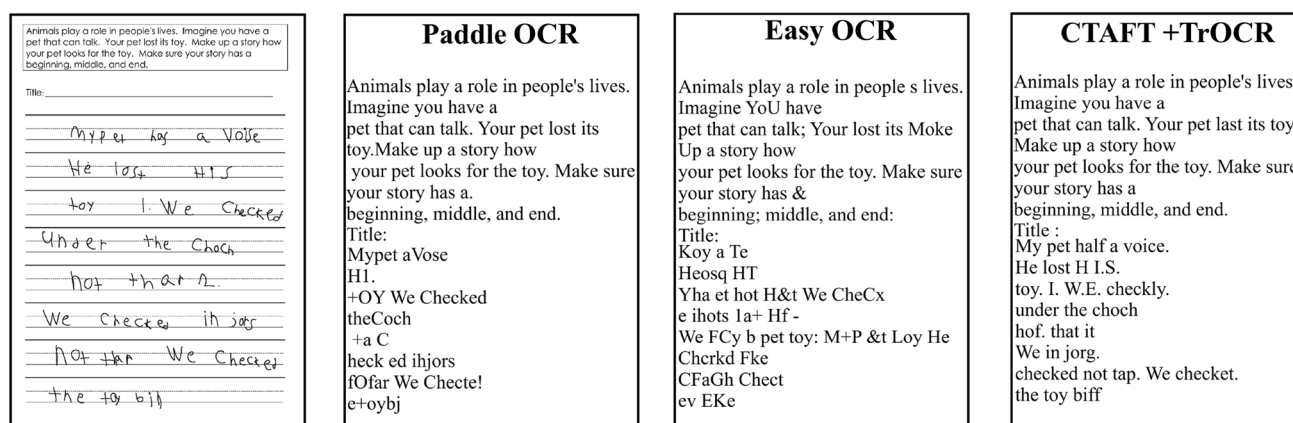


Fig. 9 Recognition of online handwriting sample using PaddleOCR, EasyOCR, and hybrid method CRAFT with TrOCR

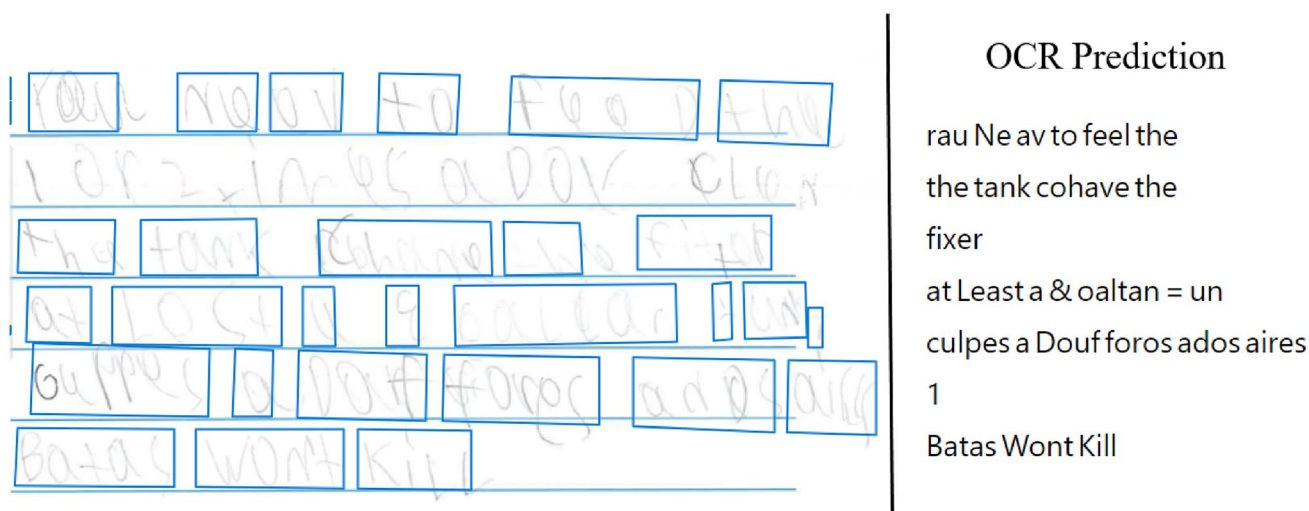


Fig. 10 Recognition of an offline handwriting sample using Microsoft Azure OCR

recognition accuracy to facilitate reliable and accurate recognition at the level necessary for our screener framework.

Semantic Analysis (SemA)

The Semantic Analysis (SemA) module will use Natural Language Processing (NLP) techniques to analyze the text data obtained from the Handwriting Recognition (HWR) module. It explores the context and meaning of the recognized text, offering a level of analysis beyond basic character recognition. This can be useful in the identification of behavioral indicators such as letter *transpositions*, *spelling errors*, *reversals*, *inversions*, and *abandoned words*. To address this, the SemA module will utilize tailored models that will be developed and trained on a curated dataset that includes labeled samples reflecting these characteristics. Data augmentation techniques also could be employed

to enhance the dataset, thereby improving the model's capability to accurately detect these errors across different handwriting styles and conditions.

As mentioned in the previous section, to accurately identify certain behavioral indicators such as letter reversals and inversions, the Handwriting Recognition (HWR) and Semantic Analysis (SemA) modules need to operate collaboratively. The HWR module transcribes the textual content from a child's handwriting sample, while the SemA module provides deeper contextual analysis to detect and interpret errors like spelling mistakes, reflections, and reversals. Understanding the semantic meaning of what the child intended to write is critical in analyzing such errors effectively. For instance in Fig. 11, what the child actually wrote, from an OCR transcription perspective, is: "The **9lant** thatI discovered looks like poop so I'm **eona** call it poop and **decause** it has poop". Here, the words "9lant", "eona",

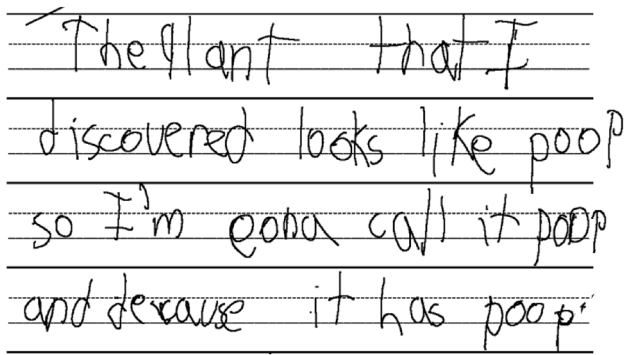


Fig. 11 A handwriting sample from first grade

and “decause” are not only spelling errors but also contain reversed letters. Specifically, “9lant” is intended to be “plant”, where the “9” is a reversed “p”; “eona” is meant to be “gonna”, where the “e” is a reversed “g”; and “decause” is intended as “because”, where the “d” is reversed “b”. These errors cannot be fully identified by HWR alone, as they require a semantic understanding of the sentence to interpret the child’s intention and resolve ambiguities. To address these types of errors, the SemA module must analyze the syntactic and semantic structure of the sentence, comparing it to plausible word choices and sentence formations.

Transpositions occur when letters within a word are in the incorrect order, while *abandoned words* are instances where writing is incomplete. Segmenting words from the handwriting samples and using a sequence-to-sequence model or a recurrent neural network (RNN) with attention mechanisms

could be strategies to address these issues. These models will be trained on extensive datasets to recognize the typical order of letters within words. The model can then analyze the sequence of letters in each word and detect anomalies that indicate transpositions. An NLP model that understands context and sentence structure will be required for identifying abandoned words. This model, trained on a large corpus of text, will discern word usage and placement within sentences, using contextual clues to determine if a word is unintentionally abandoned or ends abruptly. Additionally, integrating spell checkers and grammar tools can also aid in identifying *spelling errors* within the words.

Some OCR systems may automatically correct errors by predicting the intended word. A prime example of this is TrOCR, which integrates a language model based on RoBERTa (a robustly optimized BERT approach). This integration allows TrOCR to perform automatic corrections by understanding the context around the words. However, this automatic correction can obfuscate the system’s ability to detect the original errors, failing to accurately identify specific mistakes. For instance, Fig. 12 illustrates a situation in which the OCR system automatically rectifies a spelling error made by a child, as indicated by the blue rectangles, where “wen” is corrected to “when.” On the other hand, in Fig. 13, the OCR system accurately reproduces the child’s spelling error. Furthermore, some OCR systems can occasionally be seen to misidentify the case or correct the case to follow grammatical conventions in generating the output. For example, the OCR corrected the letter “B” in the word “ScruB” to lowercase in Fig. 13. This instability in

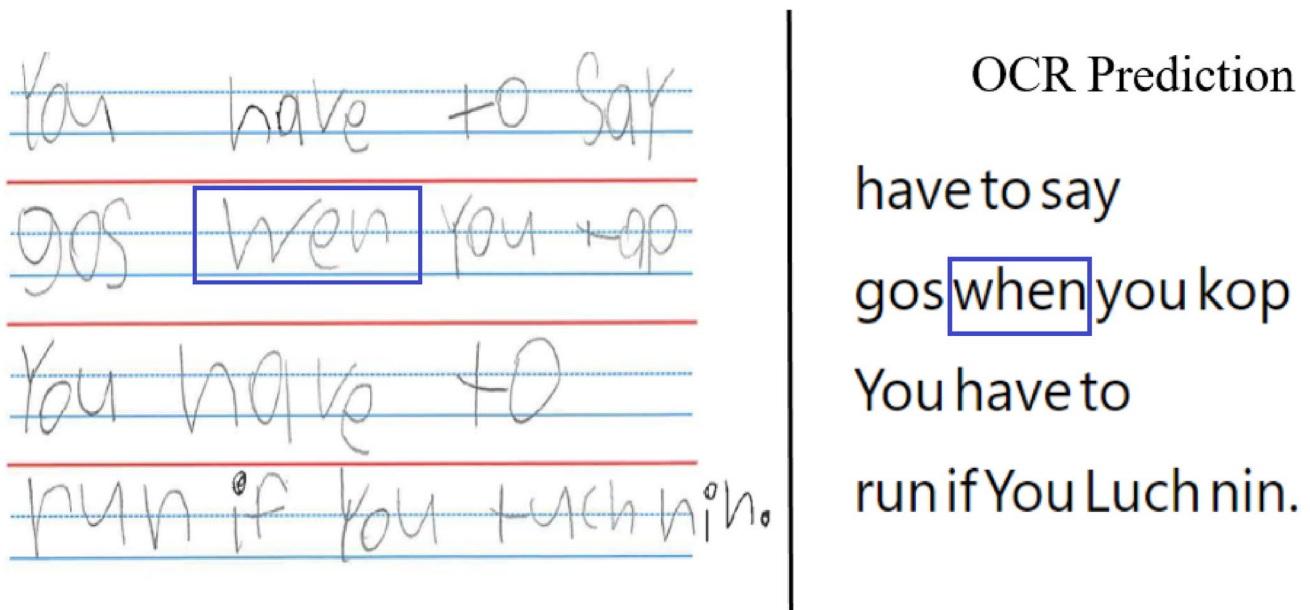


Fig. 12 Challenges in semantic analysis: auto correction of spelling errors

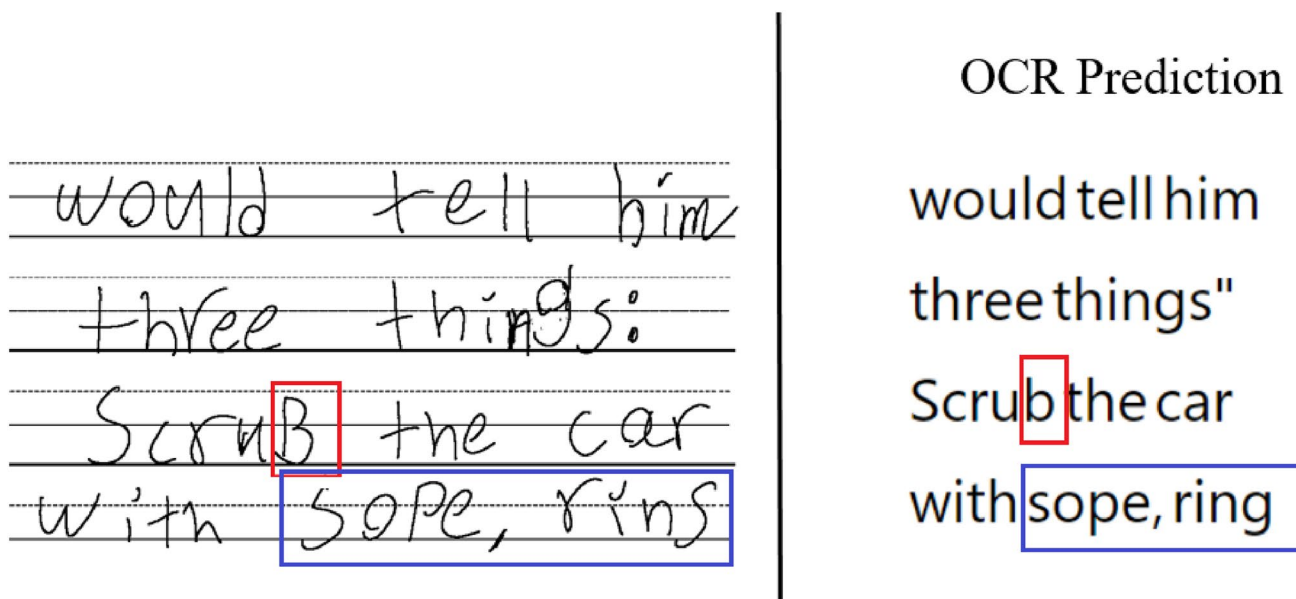


Fig. 13 Challenges in semantic analysis: difficulty in detecting the occurrence of case inconsistencies and identifying spelling errors within the text

prediction accuracy in off-the-shelf OCR systems can have a substantial effect on the reliability of semantic analysis.

The SemA module will also need to analyze sentence structures and word usage to detect patterns of repetition and evaluate complex concepts such as coherence, the organization of ideas within the text, the writer's capacity to logically structure their thoughts, and the quality and extent of the vocabulary used. The module has to effectively and reliably identify areas of concern in a child's linguistic development, such as a limited vocabulary, deficient concepts, or poor organization of ideas, through these analyses. This offers exciting research for the field.

In summary, the SemA module will not only need to leverage state-of-the-art NLP techniques but also significantly improve upon them to develop the capabilities needed for the screener framework.

Post-processing and Fusion

The Post-processing and Fusion module integrates insights from other modules to form a comprehensive evaluation of the handwriting sample, ultimately leading to a more accurate screening of potential learning disabilities. This module consolidates the insights from the Temporal and Other Analysis (TA), Structural Analysis (StA), Handwriting Recognition (HWR), and Semantic Analysis (SemA) modules to detect the 17 behavioral indicators defined in the Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC). By correlating the outputs from these modules, the Post-processing and Fusion

module ensures that the screener effectively identifies potential indicators of dyslexia and dysgraphia, offering an early screening tool for practitioners.

The module begins by aggregating the outputs from each analytical component. The TA module provides dynamic insights into handwriting, such as writing speed, pressure, and pen trajectory, which help identify motor challenges during the writing process. The StA module evaluates physical features of handwriting, including letter alignment, spacing, and margin use, which are critical for identifying structural irregularities. The HWR module converts handwritten samples into digital text, while the SemA module evaluates the semantic and syntactic meaning of the transcribed content. By combining insights from both modules, the system identifies higher-level indicators such as spelling errors, transpositions, abandoned words, letter reversals, and inversions.

The fusion process involves aligning these diverse outputs to form a multidimensional assessment of the handwriting sample. Each module's contribution is mapped to specific behavioral indicators in the DDBIC. For instance, temporal variations in writing speed identified by the TA module are cross-referenced with structural inconsistencies, such as atypical spacing or alignment, from the StA module. Similarly, spelling errors and letter reversals detected by the HWR and SemA modules are evaluated in the context of the writing's overall structure and semantics. This integrated approach enables the framework to detect subtle yet critical patterns that align with the DDBIC indicators.

By leveraging this multidimensional analysis, the Post-processing and Fusion module provides a holistic view of the handwriting sample, ensuring that all relevant behavioral indicators are captured. This module combines data from various dimensions, including temporal, structural, textual, and semantic aspects, resulting in a comprehensive and scalable tool for early screening. Importantly, the focus is not on correcting the spelling or grammatical in the handwriting sample but on detecting and contextualizing the specific features that align with the 17 DDBIC indicators, ensuring the screener is precisely tailored for its intended purpose.

In summary, the Post-processing and Fusion module transforms raw data from handwriting samples into actionable insights, bridging the gap between machine-generated outputs and practitioner needs. By integrating the strengths of all modules, this component acts as the backbone of the AI-enhanced screener, offering a reliable, scalable, and explainable solution for early detection of dyslexia and dysgraphia. It detects and aligns specific behavioral indicators making it an invaluable tool for supporting practitioners in delivering timely interventions.

Conclusion

Dyslexia and dysgraphia are two specific learning disabilities (SLDs) that can significantly impact a child's academic and social-emotional development. Therefore, practitioners such as speech-language pathologists (SLPs) and occupational therapists play a crucial role in identifying these learning disabilities early to provide the necessary intervention support and services for a child's success. However, SLPs face numerous challenges that hinder early screening and intervention. Judiciously integrating AI technologies into the screening process and developing explainable and actionable tools could enhance efficiency and simplify tasks for SLPs, while also improving the robustness of screening. This paper surveys the domain and proposes a data collection methodology and framework for developing an AI-integrated screening tool designed to identify writing-based behavioral indicators of dyslexia and dysgraphia in children's handwriting. Specifically, this framework is discussed in the context of the Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC), a writing-based screener that includes various behavioral indicators associated with dyslexia and dysgraphia.

Specific learning disabilities in children, such as dyslexia and dysgraphia, are a significant educational concern. To provide timely intervention and support, it is essential to effectively screen for these disabilities in children. As practitioners experience many challenges, the integration of Artificial Intelligence (AI) into the screening process can not only improve its efficacy but also simplify and fortify

the approach for speech-language pathologists (SLPs), thereby enhancing its robustness. To this end, this work proposes a framework for automating screening tools like the Dysgraphia and Dyslexia Behavioral Indicator Checklist (DDBIC) using AI. By leveraging the strengths of both DDBIC and AI, screening for dyslexia and dysgraphia can become more precise and efficient, paving the way for better educational outcomes. Such advancements ensure that children with learning disabilities receive the recognition and support they need early in their academic journeys, fostering an environment where all children can succeed despite their learning challenges. Integrating technology in educational diagnostics is a promising step toward a future where learning disabilities are no longer barriers to academic and personal development.

Acknowledgements This material is based upon work supported under the AI Research Institutes program by the U.S. National Science Foundation and the Institute of Education Sciences, U.S. Department of Education, through Award # 2229873—National AI Institute for Exceptional Education. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, the Institute of Education Sciences, or the U.S. Department of Education.

Author Contributions All authors contributed to the study conception and design. The first draft of the manuscript was written by Sahana Rangasrinivasan and M. S. Sumi Suresh, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding U.S. National Science Foundation Award # 2229873.

Data Availability Not applicable.

Code Availability Not applicable.

Materials Availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Consent to Publish The authors affirm that human research participants provided informed consent for publication of the images in Figs. 5, 6 and 7.

Ethical Approval The writing samples in the manuscript were collected under a study protocol approved by the University of Nevada Reno's Institutional Review Board.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in

the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Individuals with Disabilities Education Act (IDEA). <https://sites.ed.gov/idea/>. Accessed on 06/16/2024.
2. Education Statistics NC. Students with disabilities. <https://nces.ed.gov/programs/coe/indicator/cgg>. Condition of Education. U.S. Department of Education, Institute of Education Sciences. Retrieved May 27, 2024 (2023).
3. Definition of Dyslexia-International Dyslexia Association. <https://dyslexiaida.org/definition-of-dyslexia/>. Accessed on 06/10/2024.
4. Baggett M, Diamond LL, Olszewski A. Dysgraphia and dyslexia indicators: Analyzing children's writing. *Intervention in School and Clinic*; 2023;10534512231189449.
5. Hayes AM, Dombrowski E, Shefcyk AH, Bulat J. Learning disabilities screening and evaluation guide for low-and middle-income countries; 2018.
6. State of Dyslexia | National Center on Improving Literacy. <https://improvingliteracy.org/state-of-dyslexia/index.html>. Accessed on 07/10/2024.
7. Academic Screening Tools Chart Selector | National Center on Improving Literacy. <https://improvingliteracy.org/screening-selector/index.html>. Accessed on 07/11/2024.
8. Cubed Validity. <https://www.languagedynamicsgroup.com/cubed/cubed-validity/>. Accessed on 07/11/2024.
9. Lexia RAPID Assessment. <https://charts.intensiveintervention.org/screening/tool/?id=76ccb8c5478d6b22>. Accessed on 07/10/2024.
10. Phonological Awareness Literacy Screening (PALS). <https://charts.intensiveintervention.org/screening/tool/?id=5f7db477497c846b>. Accessed on 07/10/2024.
11. Chung PJ, Patel DR, Nizami I. Disorder of written expression and dysgraphia: definition, diagnosis, and management. *Transl Pediatr*. 2020;9(Suppl 1):46.
12. Spelling Administration and Scoring Guide. <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/aimswb/aimswbplus-spelling-guide.pdf>. Accessed on 07/11/2024.
13. Pearson's Dyslexia Toolkit: Dysgraphia. <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/dyslexia/dyslexia-toolkit-dysgraphia.pdf>. Accessed on 07/12/2024.
14. Beery KE. Developmental test of visual-motor integration: administration, scoring and teaching manual. London: Modern Curriculum Press; 1989.
15. Barnett AL, Henderson SE, Scheib B, Schulz J. Development and standardization of a new handwriting speed test: the detailed assessment of speed of handwriting. In: BJEP Monograph Series II, Number 6-Teaching and Learning Writing, vol. 137. London: British Psychological Society; 2009. p. 137–57.
16. Feifer S. Feifer assessment of writing. London: PAR, Inc.; 2020.
17. Treatment and Testing for Kids with Learning Disabilities. <https://www.lexercise.com/>. Accessed on 07/11/2024.
18. Henderson SE, Sugden D, Barnett AL. Movement assessment battery for children-2. Research in developmental disabilities; 1992.
19. Free Handwriting Proficiency Screener-Learning Without Tears®. <https://www.lwtears.com/resources/screener-handwriting-proficiency#:~:text=TheThe%20Screener%20of%20Handwriting%20Proficiency&text=Administer%20the%20Screener%20to%20your.you%20better%20focus%20handwriting%20instruction>. Accessed on 07/11/2024.
20. Breaux KC. WIAT 4: Wechsler individual achievement test. London: Pearson; 2020.
21. Spoon K, Crandall D, Siek K. Towards detecting dyslexia in children's handwriting using neural networks. In: Proceedings of the International Conference on Machine Learning AI for Social Good Workshop, Long Beach, CA, USA; 2019. p. 1–5.
22. Rosli MSAB, Isa IS, Ramlan SA, Sulaiman SN, Marzuki MIF. Development of CNN transfer learning for dyslexia handwriting recognition. In: 2021 11th IEEE international conference on control system, computing and engineering (ICCSCCE); 2021. p. 194–9. <https://doi.org/10.1109/ICCSCCE52189.2021.9530971>.
23. Dyslexia Handwriting Dataset. <https://www.kaggle.com/datasets/drizasazanitaisa/dyslexia-handwriting-dataset>. Accessed on 07/10/2024.
24. Alkhurayyif Y, Sait ARW. Deep learning-based model for detecting dyslexia using handwritten images. *J Disab Res*. 2023;2(4):89–98.
25. Alqahtani ND, Alzahrani B, Ramzan MS. Detection of dyslexia through images of handwriting using hybrid AI approach. *Int J Adv Comput Sci Appl*. 2023;14(10):1.
26. Sreekumar S, Lijiya A. Comparative study of CNN models on the classification of dyslexic handwriting. In: 2022 IEEE Bombay section signature conference (IBSSC). IEEE; 2022. p. 1–4.
27. Jasira K, Laila V. Dyslexiscan: a dyslexia detection method from handwriting using CNN LSTM model. In: 2023 international conference on innovations in engineering and technology (ICIET). IEEE; 2023. p. 1–6.
28. Isa IS, Rahimi WNS, Ramlan SA, Sulaiman SN. Automated detection of dyslexia symptom based on handwriting image for primary school children. *Proc Comput Sci*. 2019;163:440–9.
29. Isa IS, Zahir MA, Ramlan SA, Wang L-C, Sulaiman SN. CNN comparisons models on dyslexia handwriting classification. *ESTEEM Acad J*. 2021;17:12–25.
30. Kariyawasam R, Nadeeshani M, Hamid T, Subasinghe I, Samarasinghe P, Ratnayake P. Pubudu: deep learning based screening and intervention of dyslexia, dysgraphia and dyscalculia. In: 2019 14th conference on industrial and information systems (ICIIS). IEEE; 2019. p. 476–81.
31. Devi A, Kavya G. Dysgraphia disorder forecasting and classification technique using intelligent deep learning approaches. *Prog Neuropsychopharmacol Biol Psychiatry*. 2023;120: 110647.
32. Hewapathirana C, Abeysinghe K, Maheshani P, Liyanage P, Krishara J, Thelijagoda S. A mobile-based screening and refinement system to identify the risk of dyscalculia and dysgraphia learning disabilities in primary school students. In: 2021 10th international conference on information and automation for sustainability (ICIAfS). IEEE; 2021. p. 287–92.
33. Kunhoth J, Al Maadeed S, Saleh M, Akbari Y. CNN feature and classifier fusion on novel transformed image dataset for dysgraphia diagnosis in children. *Expert Syst Appl*. 2023;231: 120740.
34. Kunhoth J, Al Maadeed S, Saleh M, Akbari Y. Exploration and analysis of on-surface and in-air handwriting attributes to improve dysgraphia disorder diagnosis in children based on machine learning methods. *Biomed Signal Process Control*. 2023;83: 104715.
35. Drotár P, Dobeš M. Dysgraphia detection through machine learning. *Sci Rep*. 2020;10(1):21541.
36. Mekyska J, Faundez-Zanuy M, Mzourek Z, Galaz Z, Smekal Z, Rosenblum S. Identification and rating of developmental dysgraphia by handwriting analysis. *IEEE Trans Hum Mach Syst*. 2016;47(2):235–48.
37. Asselborn T, Gargot T, Kidziński Ł, Johal W, Cohen D, Jolly C, Dillenbourg P. Automated human-level diagnosis of dysgraphia using a consumer tablet. *NPJ Dig Med*. 2018;1(1):42.

38. Devillaine L, Lambert R, Boutet J, Aloui S, Brault V, Jolly C, Labyt E. Analysis of graphomotor tests with machine learning algorithms for an early and universal pre-diagnosis of dysgraphia. *Sensors*. 2021;21(21):7026.
39. Dankovičová Z, Hurtuk J, Fecil'ak P. Evaluation of digitalized handwriting for dysgraphia detection using random forest classification method. In: 2019 IEEE 17th international symposium on intelligent systems and informatics (SISY). IEEE; 2019. p. 000149–000154.
40. Deschamps L, Devillaine L, Gaffet C, Lambert R, Aloui S, Boutet J, Brault V, Labyt E, Jolly C. Development of a pre-diagnosis tool based on machine learning algorithms on the BHK test to improve the diagnosis of dysgraphia. *Adv Artif Intel Mach Learn*. 2021;1(2):114–35.
41. Rosenblum S, Dror G. Identifying developmental dysgraphia characteristics utilizing handwriting classification methods. *IEEE Trans Hum Mach Syst*. 2016;47(2):293–8.
42. Sihwi SW, Fikri K, Aziz A. Dysgraphia identification from handwriting with support vector machine method. *J Phys Conf Ser*. 2019;1201:012050.
43. Mekyska J, Galaz Z, Safarova K, Zvoncak V, Mucha J, Smekal Z, Ondrackova A, Urbanek T, Havigerová JM, Bednarova J. Computerised assessment of graphomotor difficulties in a cohort of school-aged children. In: 2019 11th international congress on ultra modern telecommunications and control systems and workshops (ICUMT). IEEE; 2019. p. 1–6.
44. Zvoncak V, Mekyska J, Safarova K, Smekal Z, Brezany P. New approach of dysgraphic handwriting analysis based on the tunable q-factor wavelet transform. In: 2019 42nd international convention on information and communication technology, electronics and microelectronics (MIPRO). IEEE; 2019. p. 289–294.
45. Zvoncak V, Mekyska J, Safarova K, Galaz Z, Mucha J, Kiska T, Smekal Z, Losenicka B, Cechova B, Francova P. Effect of stroke-level intra-writer normalization on computerized assessment of developmental dysgraphia. In: 2018 10th international congress on ultra modern telecommunications and control systems and workshops (ICUMT). IEEE; 2018. p. 1–5.
46. Dimauro G, Bevilacqua V, Colizzi L, Di Piero D. Testgraphia, a software system for the early diagnosis of dysgraphia. *IEEE Access*. 2020;8:19564–75.
47. Ghouse F, Paranjothi K, Vaithyanathan R. Dysgraphia classification based on the non-discrimination regularization in rotational region convolutional neural network. *Int J Intell Eng Syst*. 2022;15(1):1.
48. Kunthoth J, Al-Maadeed S, Kunthoth S, Akbari Y, Saleh M. Automated systems for diagnosis of dysgraphia in children: a survey and novel framework. *Int J Doc Anal Recogn (IJ DAR)*. 2024;2024:1–29.
49. Yogarajah P, Bhushan B. Deep learning approach to automated detection of dyslexia-dysgraphia. In: The 25th IEEE international conference on pattern recognition; 2020.
50. Rashid HA, Malik T, Siddiqui I, Bhatti N, Samad A. Dysign: towards computational screening of dyslexia and dysgraphia based on handwriting quality. In: Proceedings of the 22nd annual ACM interaction design and children conference; 2023. p. 532–6.
51. Lakretz Y, Chechik G, Friedmann N, Rosen-Zvi M. Probabilistic graphical models of dyslexia. In: Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining; 2015. p. 1919–28.
52. Xiong J, Yin H, Pan M. Application of image classification based on improved LSTM in internet reading therapy platform. *IEEE Access*. 2023;2023:1.
53. Radford J, Richard G, Richard H, Serrurier M. Detecting dyslexia from audio records: an AI approach. In: HEALTHINF; 2021. p. 58–66.
54. Richard G, Serrurier M. Dyslexia and dysgraphia prediction: a new machine learning approach. Preprint [arXiv:2005.06401](https://arxiv.org/abs/2005.06401); 2020.
55. Abdul Hamid SS, Admodisastro N, Manshor N, Kamaruddin A, Ghani AAA. Dyslexia adaptive learning model: student engagement prediction using machine learning approach. In: Recent advances on soft computing and data mining: proceedings of the third international conference on soft computing and data mining (SCDM 2018), Johor, Malaysia, February 06–07, 2018. London: Springer; 2018. p. 372–84.
56. Rello L, Ballesteros M. Detecting readers with dyslexia using machine learning with eye tracking measures. In: Proceedings of the 12th international web for all conference; 2015. p. 1–8.
57. Nilsson Benfatto M, Öqvist Seimyr G, Ygge J, Pansell T, Rydberg A, Jacobson C. Screening for dyslexia using eye tracking during reading. *PLoS ONE*. 2016;11(12):0165508.
58. Asvestopoulou T, Manousaki V, Psistakis A, Smyrnakis I, Andreiadakis V, Aslanides IM, Papadopoulou M. Dyslexml: screening tool for dyslexia using machine learning. Preprint [arXiv:1903.06274](https://arxiv.org/abs/1903.06274); 2019.
59. Shalileh S, Ignatov D, Lopukhina A, Dragoy O. Identifying dyslexia in school pupils from eye movement and demographic data using artificial intelligence. *PLoS ONE*. 2023;18(11):0292047.
60. Cui Z, Xia Z, Su M, Shu H, Gong G. Disrupted white matter connectivity underlying developmental dyslexia: a machine learning approach. *Hum Brain Mapp*. 2016;37(4):1443–58.
61. Płoński P, Gradkowski W, Altarelli I, Monzalvo K, Ermingen-Marbach M, Grande M, Heim S, Marchewka A, Bogorodzki P, Ramus F. Multi-parameter machine learning approach to the neuroanatomical basis of developmental dyslexia. *Hum Brain Mapp*. 2017;38(2):900–8.
62. Zahia S, Garcia-Zapirain B, Saralegui I, Fernandez-Ruanova B. Dyslexia detection using 3D convolutional neural networks and functional magnetic resonance imaging. *Comput Methods Programs Biomed*. 2020;197: 105726.
63. Frid A, Manevitz LM. Features and machine learning for correlating and classifying between brain areas and dyslexia. Preprint [arXiv:1812.10622](https://arxiv.org/abs/1812.10622); 2018.
64. Rezvani Z, Zare M, Žarić G, Bonte M, Tijms J, Molen M, Fraga González G. Machine learning classification of dyslexic children based on EEG local network features. *BioRxiv*. 2019;2019: 569996.
65. Pen Tablet / Digitizer Tablet. <https://www.wacom.com/en-in/products/pen-tablets>. Accessed on 06/30/2024.
66. Gemelli A, Marinai S, Vivoli E, Zappaterra T. Deep-learning for dysgraphia detection in children handwritings. In: Proceedings of the ACM symposium on document engineering 2023; 2023. p. 1–4.
67. Asselborn T, Chapatte M, Dillenbourg P. Extending the spectrum of dysgraphia: a data driven strategy to estimate handwriting quality. *Sci Rep*. 2020;10(1):3140.
68. Olive T. Working memory in writing: empirical evidence from the dual-task technique. *Eur Psychol*. 2004;9(1):32–42.
69. Berninger VW, Richards TL, Abbott RD. Differential diagnosis of dysgraphia, dyslexia, and OWL LD: behavioral and neuroimaging evidence. *Read Writ*. 2015;28:1119–53.
70. Döhla D, Heim S. Developmental dyslexia and dysgraphia: What can we learn from the one about the other? *Front Psychol*. 2016;6:2045.
71. Winkes J. Isolierte rechtschreibstörung: eigenständiges störungsbild oder leichte form der lese-rechtschreibstörung? eine untersuchung der kognitiv-linguistischen informationsverarbeitungs-kompetenzen von kindern mit schriftspracherwerbsstörungen: Julia winkes; 2014.

72. Marti U-V, Bunke H. The IAM-database: an English sentence database for offline handwriting recognition. *Int J Doc Anal Recogn.* 2002;5:39–46.
73. Augustin E, Carré M, Grosicki E, Brodin J-M, Geoffrois E, Prêteux F. Rimes evaluation campaign for handwritten mail processing. In: *International workshop on frontiers in handwriting recognition (IWFHR'06)*; 2006. p. 231–5.
74. Kleber F, Fiel S, Diem M, Sablatnig R. Cvl-database: an off-line database for writer retrieval, writer identification and word spotting. In: *2013 12th international conference on document analysis and recognition*. IEEE; 2013. p. 560–4.
75. Common Core State Standards Initiative: K-12 English Language Arts Standards. Accessed: 2024-05-24 (2010). https://webapp-strap1-paas-prod-nde-001.azurewebsites.net/uploads/K_12_ELA_Standards_ADA_Accessible_ee73596bc7.pdf.
76. Papavassiliou V, Stafylakis T, Katsouros V, Carayannis G. Handwritten document image segmentation into text lines and words. *Pattern Recogn.* 2010;43(1):369–77.
77. Louloudis G, Gatos B, Pratikakis I, Halatsis C. Text line detection in handwritten documents. *Pattern Recogn.* 2008;41(12):3758–72.
78. Razak Z, Zulkiflee K, Idris MYI, Tamil EM, Noor MNM, Salleh R, Yaakob M, Yusof ZM, Yaacob M. Off-line handwriting text line segmentation: a review. *Int J Comput Sci Netw Secur.* 2008;8(7):12–20.
79. Baek Y, Lee B, Han D, Yun S, Lee H. Character region awareness for text detection. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*; 2019. p. 9365–9374.
80. Zhou X, Yao C, Wen H, Wang Y, Zhou S, He W, Liang J. East: an efficient and accurate scene text detector. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2017. p. 5551–60.
81. Lu Y, Shridhar M. Character segmentation in handwritten words—an overview. *Pattern Recogn.* 1996;29(1):77–96.
82. Casey RG, Lecolinet E. A survey of methods and strategies in character segmentation. *IEEE Trans Pattern Anal Mach Intell.* 1996;18(7):690–706.
83. Rajput V, Jayanthi N, Indu S. An efficient character segmentation algorithm for connected handwritten documents. In: *Document analysis and recognition: 4th workshop, DAR 2018, held in conjunction with ICVGIP 2018, Hyderabad, India, December 18, 2018, Revised Selected Papers 4*. London: Springer; 2019. p. 97–105.
84. Microsoft: Azure AI Vision Services. 2023. <https://azure.microsoft.com/en-us/products/ai-services/ai-vision>. Accessed: 2024-06-15.
85. Services AW. Amazon Textract. <https://aws.amazon.com/textract/>. Accessed: 2024-11-05.
86. Du Y, Li C, Guo R, Yin X, Liu W, Zhou J, Bai Y, Yu Z, Yang Y, Dang Q, et al. Pp-ocr: a practical ultra lightweight OCR system; 2020. Preprint [arXiv:2009.09941](https://arxiv.org/abs/2009.09941); 2009.
87. JaidedAI: EasyOcr. GitHub repository; 2022. <https://github.com/JaidedAI/EasyOCR>.
88. Smith R. An overview of the tesseract OCR engine. In: *ICDAR'07: proceedings of the ninth international conference on document analysis and recognition*. IEEE Computer Society, Washington, DC, USA; 2007. p. 629–33. <https://storage.googleapis.com/pub-tools-public-publication-data/pdf/33418.pdf>.
89. Li M, Lv T, Chen J, Cui L, Lu Y, Florencio D, Zhang C, Li Z, Wei F. Trocr: transformer-based optical character recognition with pre-trained models. In: *Proceedings of the AAAI conference on artificial intelligence*, vol. 37; 2023. p. 13094–102.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.