DOI: https://doi.org/10.4038/jdrra.v2i1.56

The Journal of Desk Research Review and Analysis, Vol. 2, Issue 1, 2024, 216-234

ADVANCED COMPUTATIONAL TECHNIQUES FOR DYSGRAPHIA PREDICTION THROUGH HANDWRITING RECOGNITION USING MACHINE LEARNING AND DEEP LEARNING METHODS

S Weraduwa[1], PPGD Asanka[2] and TV Mahanama[3]

Abstract

This systematic review critically examines the utilization of machine learning (ML) and deep learning (DL) techniques in detecting and predicting Dysgraphia through handwriting recognition. The primary research objective is to systematically assess the significance, contributions, performance, and limitations of recent technological advances in identifying dysgraphic patients. Employing a structured search and selection methodology based on the PRISMA protocol, 34 peer-reviewed studies from key academic databases, including IEEE Xplore, PubMed, Scopus, Google Scholar, and Science Direct, were comprehensively analyzed. The review methodology focused extensively on data collection methods, ethical considerations, preprocessing approaches, feature extraction and selection techniques, and the performance metrics of various ML models in dysgraphia prediction. Our key findings indicate that a Hybrid AI method combining convolutional neural networks and support vector machines demonstrates high accuracy rates, reaching up to 99.33%. This systematic review compares machine learning and deep learning use in diagnosing Dysgraphia through handwriting recognition with traditional cognitive evaluations, highlighting the latter's lack of scalability and high labour demands. Additionally, this review highlights that research on Dysgraphia has mainly focused on Latin scripts, overlooking the unique challenges of non-Latin scripts. It stresses the importance of including diverse linguistic data for more accurate global diagnoses. Additionally, Dysgraphia is influenced by cognitive, motor, and language skills, underscoring the need for multimodal data to enhance early detection and treatment effectiveness. The paper advocates for research that integrates various data types and expands linguistic diversity in training datasets, aiming to improve the diagnostic accuracy and global applicability of dysgraphia detection tools.

Keywords: Deep learning, Dysgraphia, Dyslexia, Handwriting Recognition

---

[1] Department of Industrial Management, University of Kelaniya, Sri Lanka

Email: sulaksh-im19112@stu.kln.ac.lk https://orcid.org/0009-0009-1937-0179

[2] Senior Lecturer, Department of Industrial Management, University of Kelaniya, Sri Lanka

Email: dasanka@kln.ac.lk https://orcid.org/0000-0002-3433-5533

[3] Lecturer, Department of Industrial Management, University of Kelaniya, Sri Lanka

Email: thilinim@kln.ac.lk https://orcid.org/0000-0003-0536-0040

Received date: 09.07.2024
Accepted date: 22.10.2024

Introduction

Learning disabilities encompass a range of challenges that affect skill development but do not imply intellectual deficiency. These disabilities often undermine children's self-esteem and confidence, particularly in reading, writing, and math. Dysgraphia is a specific learning disability that primarily affects written expression, including spelling, grammar, and handwriting. It is characterized by messy handwriting, irregular letter spacing and capitalization, discomfort during writing, and difficulties with fine motor skills and spelling. Dysgraphia may also coexist with other disorders, such as autism spectrum disorder (ASD), Dyslexia, or attention deficit hyperactivity disorder (ADHD), highlighting the necessity for prompt diagnosis and treatment (Lopez et al., 2018).

A comprehensive evaluation involving child psychiatrists and neuropediatric specialists, along with the collaboration of psychometricians and speech therapists, is essential for an accurate assessment (Kunhoth et al., 2023a) of Dysgraphia. Dysgraphia symptoms are detected using conventional methods, such as various screening tests. This traditional approach has several limitations, including being time intensive as assessors need to observe meticulously and record behaviours indicative of Dysgraphia. The process is also demanding as it requires a dysgraphia specialist to manage the test actively, particularly because dyslexic children may quickly lose interest and struggle to focus during the evaluation. Furthermore, this method often results in prolonged periods to obtain final results. However, early identification of Dysgraphia is crucial to providing the necessary support and enabling therapeutic interventions, enhancing children's ability to cope with educational and social demands (Asselborn et al., 2018; Tao & Rapp, 2019).

Despite the promising advances in artificial intelligence (AI), significant challenges hinder the effective translation of handwriting analysis into reliable diagnostic outcomes for Dysgraphia. The complexity of handwriting as a data type, coupled with the diverse manifestations of Dysgraphia and its frequent co-occurrence with other learning disabilities such as ADHD and Dyslexia, complicates the development of universally effective diagnostic tools. These complications are exacerbated by the variable nature of dysgraphia symptoms and the impact of cognitive, motor, and linguistic factors inherent in the disorder. Moreover, technological limitations in machine learning models, such as data handling and model generalization, alongside methodological constraints, such as sample diversity and ethical concerns, pose additional barriers. This research aims to address the following questions by its conclusion:

How effective are current machine learning and deep learning approaches in accurately predicting Dysgraphia through handwriting analysis?

Q2- What are the main technological and methodological limitations faced by these advanced diagnostic tools?

Q3- Beyond solely using handwriting to detect Dysgraphia, are there any studies that employ multimodal approaches by integrating cognitive and motor skills along with handwriting to identify Dysgraphia?

By the end of this research, the following objectives are expected to be accomplished.

RO1-**To Evaluate the Current Capabilities of ML and DL in Dysgraphia Diagnosis:** Assess how machine learning and deep learning technologies can predict Dysgraphia through handwriting analysis, considering the computational methods' accuracy.

RO2**- To Identify Technological and Methodological Limitations:** Explore significant limitations within these advanced diagnostic tools, including data diversity, model overfitting, and the generalization capabilities of the models. Emphasize the need to enhance linguistic diversity in training datasets to improve global applicability and diagnostic accuracy.

RO3- **To introduce a multimodal approach for dysgraphia detection**: This approach will integrate cognitive, motor, and language assessments to enhance the accuracy of machine learning models in diagnosing children with Dysgraphia.

This review examines explicitly the processes involved in data collection, ethical considerations, preparation and preprocessing, feature extraction and selection, training and classification, performance evaluation and recommendations. ML, an expanding area of research, develops predictive models from specific datasets. This review compares how traditional machine and deep learning approaches have progressed in predicting Dysgraphia using handwriting recognition.

Current Approaches to Dysgraphia Detection

The literature discusses leveraging ML and DL techniques for diagnosing and intervening in learning disabilities among children. Notably, (Hewapathirana et al., 2021) and (Kariyawasam et al., 2019) introduced "Nana Shilpa" and "Pubudu," mobile applications that utilize convolutional neural networks (CNN) and other machine learning models to detect and address Dyscalculia, Dysgraphia, and Dyslexia in Sri Lankan children. "Nana Shilpa" employs CNNs and SVM, demonstrating remarkable accuracies of 92% for Dyscalculia and 99% for Dysgraphia. Also, "Pubudu" integrates voice recognition and employs a multisensory approach, achieving accuracies of 88% for Dyslexia, 58% for Dysgraphia, and 99% for Dyscalculia. Despite their high efficacy, both applications have challenges, such as limited language support and a general lack of awareness about learning disabilities. "Pubudu" (Kariyawasam et al., 2019) is noted for its ongoing efforts to enhance model accuracy and expand linguistic capabilities.

Further studies continue to exploit deep learning's capabilities. The study (Vaidya et al., 2022) applies MediaPipe Hands technology and machine learning for air writing recognition that assists dyslexic learners and achieves substantial accuracy. However, the complexity of setup and the need for further accuracy enhancement remain significant challenges. Another research (Kunhoth et al., 2023b) explores the use of multiple machine learning classifiers, including the k-nearest neighbours algorithm (KNN), SVM, and AdaBoost, for diagnosing Dysgraphia. The AdaBoost classifier reports an impressive accuracy of 80.8%, though the study acknowledges the limitation of small dataset sizes, which may impede generalization.

Integrating CNN with recurrent neural networks (Doshi et al., 2023) identifies risks of Dysgraphia from handwriting errors with a 73% accuracy rate. In their recent study, (Anand et al., 2023) enhanced CNN's performance by augmenting data samples, achieving an 84% accuracy rate. Despite these advances, both studies face challenges related to data diversity and potential model overfitting. Similarly, (Sharmila et al., 2023; Yogarajah, n.d.) apply CNN to dyslexia-dysgraphia feature detection, achieving up to 97.98% accuracy. However, these studies also highlight significant challenges, such as high diagnosis costs and the risk of data overfitting. Collectively, these studies have the potential and limitations of applying CNN and other advanced analytics in the diagnostic processes of learning disabilities, advocating for ongoing improvement in technology accuracy and practical application.

Recent studies have progressively harnessed the power of ensemble learning and sophisticated machine learning techniques to enhance the detection and diagnostic accuracy of learning disabilities such as Dyslexia and Dysgraphia. The study by (Gasmi et al., 2024) employs a hybrid ML methodology that incorporates genetic algorithms to optimize ensemble classifiers, achieving.

A notable 91% accuracy in dyslexia identification is found in a related study (AlGhamdi, 2022), which addresses the challenge of class imbalance by implementing techniques such as AdaBoost and RUSBoost alongside a KNN ensemble, reaching accuracies as high as 99.9%. Additionally, (Kunhoth et al., 2023a) integrate transfer learning with SVM, AdaBoost, and random forest to diagnose Dysgraphia with a high accuracy of 97.3%. These studies provide examples of the effectiveness of combining ensemble methods with machine learning to refine diagnostic precision, though they acknowledge existing challenges such as class imbalance and data quality.

The use of CNN and long short-term memory (LSTM) networks has demonstrated considerable promise in tackling learning disabilities. Utilizing CNNs for extracting features from handwriting images and combining them with LSTMs for sequential analysis, (KT et al., 2023) achieve significant accuracy, though they encounter issues such as small sample sizes and potential model overfitting. Extending this approach, (Liu et al., 2024) incorporate positional encoding and self-attention mechanisms into the CNN-LSTM framework, specifically targeting early dyslexia detection in Chinese-speaking children and reaching up to 85% accuracy. However, this model's complexity hinders its interpretability. Further advancing this technology, (Masood et al., 2023) combine the CNN-LSTM output with a random forest classifier, achieving impressive accuracies of over 96% and recommend further enhancements through transfer learning and more varied datasets.

Some studies explore hybrid deep-learning models to tackle various challenges associated with Dyslexia. The study (Safura et al., 2023) implements a rule-based approach combined with a personalized text simplification framework to enhance academic performance through tailored texts for dyslexic students, overcoming semantic perspective and personalization limitations. Furthermore, (Alqahtani et al., 2023) utilize a combination of CNN, SVM, and Random Forest models to predict Dyslexia from handwriting images, achieving an accuracy of 98.59%. Despite challenges such as small datasets and varying performance in tests, these studies demonstrate the efficacy of hybrid models in extracting meaningful features from complex data for accurate dyslexia detection.

In addressing the diagnosis of Dyslexia, (Walda et al., 2022) employ neural networks and ensemble techniques to evaluate cognitive skills, revealing limited predictive capacity for reading development. Another study, Alia (Alia Hussein et al., 2024), utilizes spectrogram analysis and convolutional neural networks (CNNs) to convert audio signals into visual data for early dyslexia detection, achieving high accuracy and surpassing conventional methods. While both studies highlight the efficacy of machine learning in advancing dyslexia diagnostics, they also acknowledge considerable limitations concerning the scope of research and the reliability of their models.

In tablet-based applications for dysgraphia detection, (Menikdiwela et al., n.d.) focus on analyzing healthcare and special education facilities, while (Asselborn et al., 2018) develop a diagnostic tool employing a Random Forest classifier. This latter study achieves significant efficacy, reaching 96.6% sensibility and 99.2% specificity by analyzing handwritten dynamics such as pen pressure and tilt on a consumer tablet. This approach marks a departure from traditional methods, highlighting a shift towards more objective and scalable diagnostic tools in educational technology.

Continuing with the theme of advancing the diagnosis of learning disabilities, several studies employ diverse machine-learning strategies with notable effectiveness. The study titled ("Dysgraphia Classification Based on the Non-Discrimination Regularization in Rotational Region Convolutional Neural Network" 2022) implements the NDR-R2CNN model, focusing on non-discriminative word analysis for Dysgraphia, achieving an impressive 98.2% accuracy. The study (et al., 2021) uses the Random Forest algorithm, demonstrating a high accuracy of 99.03% in dysgraphia prediction by leveraging IoT devices for enhanced data analysis. Additionally, a study (by Khan et al., 2020) develops an automated diagnostic system using various machine-learning techniques, achieving 98% accuracy in detecting Dyslexia among children. Exploring the use of linguistic features for dyslexia diagnosis in Persian-speaking children (Tolami et al., 2021) attain high precision and recall.

Finally, (Sasidhar et al., 2022) and (Ikermane and El Mouatasim, 2023) both leverage advanced ML models to enhance the detection of Dysgraphia and Dyslexia. Sasidhar et al. employ the Residual Neural Network (ResNet) combined with OpenCV for preprocessing handwriting images, achieving improved dyslexia detection rates. Furthermore (Ikermane & El Mouatasim, 2023) utilizes an Artificial Neural Network (ANN) model for dysgraphia classification, demonstrating a 96% accuracy using digital handwriting traits. These approaches underscore the potential of deep learning in automating and refining the diagnostic process, enabling earlier educational interventions.

Methodology

In this systematic review, we employed a structured search and selection process based on the PRISMA protocol to investigate a specified topic within academic literature. The process of identifying and selecting relevant articles adhered to the guidelines set forth by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2009). As depicted in Fig. 1, this process encompassed three distinct phases: identification, screening, and inclusion.

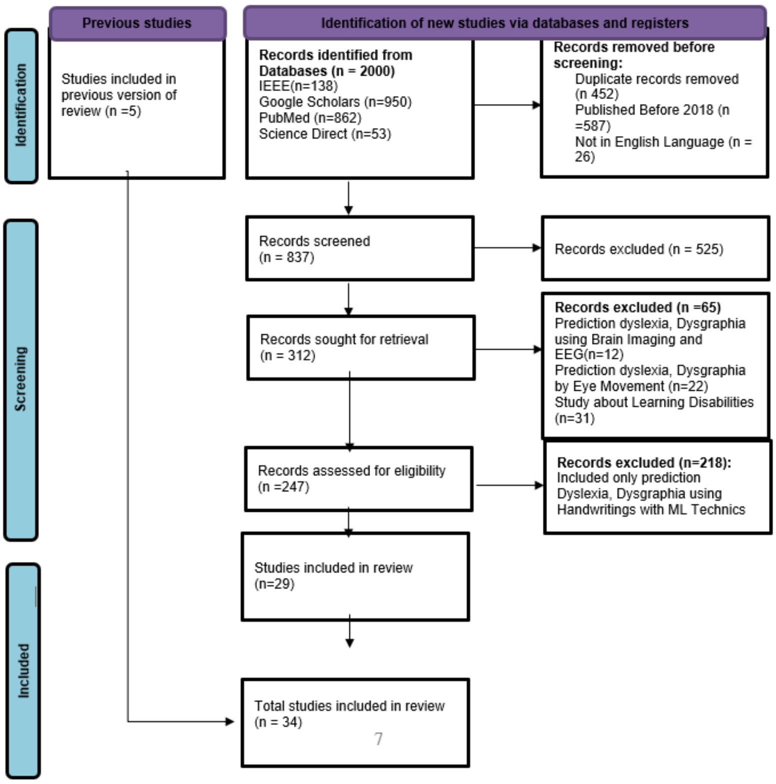


Figure 1- PRISMA Flow Diagram for Filtering 34 Articles

A multi-database search strategy was employed to conduct a comprehensive and precise literature search for the systematic. The databases selected for this search are known for their extensive coverage of fields relevant to our research objectives, such as artificial intelligence, machine learning, medical informatics, and educational technology. These databases include IEEE Xplore, PubMed, Scopus, Google Scholar, and Science Direct.

Keywords: Deep learning, Dysgraphia, Dyslexia, Handwriting Recognition

The review began with a comprehensive search of multiple databases to ensure a broad coverage of potentially relevant studies. The databases included IEEE Xplore (n=138), Google Scholar (n=950), PubMed (n=862), and Science Direct (n=53). This initial search yielded a total of 2000 records. To refine this extensive data set, we utilized Boolean operators as a fundamental search technique, allowing for nuanced and precise literature filtering. Our search strings were carefully constructed to include combinations of keywords such as "Dysgraphia," "Dyslexia" "Handwriting Recognition," "Deep Learning," and "Predictive Models," linked by Boolean connectors like AND and OR. This strategy enabled us to precisely target studies that directly addressed our research questions about deep learning applications in handwriting analysis for predicting Dysgraphia. This research search string, crafted with Boolean operators, was: ("Dysgraphia" OR "Dyslexia" AND ("Handwriting Recognition" OR "Handwriting Analysis") AND ("Deep Learning" OR "Neural Networks") AND "Predictive Models").

From the initial pool, records were first subjected to the removal of duplicates, resulting in 452 records being excluded. Additional records were excluded based on the following criteria:

- Publications before 2018 to focus on the most recent research (n=687)
- Non-English language publications to accommodate the language capabilities of the review team (n=26)

Following these exclusions, 837 records were screened for relevance based on their titles and abstracts. A focused search of these records resulted in the identification of 312 records that were considered relevant for detailed analysis. The next stage involved a thorough review of these 312 records to assess their eligibility based on more specific inclusion criteria, which focused on the study’s relevance to the primary research questions concerning the prediction of Dysgraphia.
During the full-text review, 65 number of studies were excluded for reasons including:

- Studies that solely predicted Dyslexia or Dysgraphia employing the specified technologies such as EEG or brain imaging (n=12)
- Studies focusing on eye movement analysis (n=22)
- Research related to the direct prediction or analysis of learning disabilities like Dyslexia and Dysgraphia (n= 31)
-

This review phase resulted in 247 records being selected for full-text assessment. Out of the 247 records retrieved for detailed assessment, 218 were excluded primarily because they did not meet the specific inclusion criteria of the review, which focused on dyslexia or dysgraphia prediction using handwriting with machine learning techniques.

After the detailed eligibility assessment, 29 studies met all the criteria and were included in the systematic review. These studies were then combined with five additional studies from previous reviews, bringing the total number of studies included in the final analysis to 34.

This meticulous and structured methodology enabled us to systematically explore and synthesize the most relevant and recent literature on the prediction of Dysgraphia using advanced diagnostic techniques. The review process was designed to minimize bias and maximize the comprehensiveness and relevance of the findings to the research questions posed. Table 1 shows the inclusion and exclusion criteria of articles.

Table 3-Inclusion Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
Publication Year	Articles released between 2018 and 2024 in English	Articles published before 2018
Language	Articles in English	Articles in non-English languages
Technological Focus	Articles that utilized DL techniques and ML for handwriting recognition	Articles that used only traditional Analytics and Statistical methods.
Specific Disorder Focus	Articles that focus specifically on the prediction of Dysgraphia and Dyslexia.	Articles relevant to Dyscalculia and Specific Learning Disorders.
Relevant Technologies	Articles employing handwriting analysis in conjunction with DL for dysgraphia and dyslexia prediction	Articles focusing on dysgraphia and dyslexia prediction using EEG or Brain Images and eye movement analysis

Results

Dyslexia prediction using ML and DL models consists of multiple processes, beginning with acquiring datasets and ending with evaluating the prediction models. This can be noticed in the papers chosen for review.

Data Acquisition

The initial step in predicting Dysgraphia through DL and ML involves diverse methodologies for data acquisition, focusing on a variety of handwriting data to train and evaluate models. Widely used public datasets like the Dyslexia Handwriting dataset from Kaggle and the MNIST dataset provide extensive pre-labelled samples for model training (Sasidhar et al., 2022; Vaidya et al., 2022), while custom datasets, such as those gathered using Wacom Intuos tablets, offer tailored features from specific populations (Ikermane & El Mouatasim, 2023). Sample demographics typically include children to capture developmental variations in handwriting, with dataset sizes ranging from small pilot studies to extensive collections facilitating comprehensive model training (Gemelli et al., 2023; Sharmila et al., 2023; Spoon & Crandall, n.d.; Yogarajah, n.d.). Data types extend beyond handwriting samples, incorporating linguistic essays and cognitive tests to understand the disorders more comprehensively (Tolami et al., 2021; Walda et al., 2022). Consistent data collection is ensured through standardized tests and digital tools like the BHK test and Wacom tablets, which help capture detailed handwriting features under controlled conditions (Asselborn et al., 2018; Menikdiwela et al., n.d.). This structured approach supports the robust training and validation of deep learning models essential for accurate dysgraphia prediction.

Fig. 2 shows the Data Collection Framework for Predicting Dysgraphia Using Deep Learning.

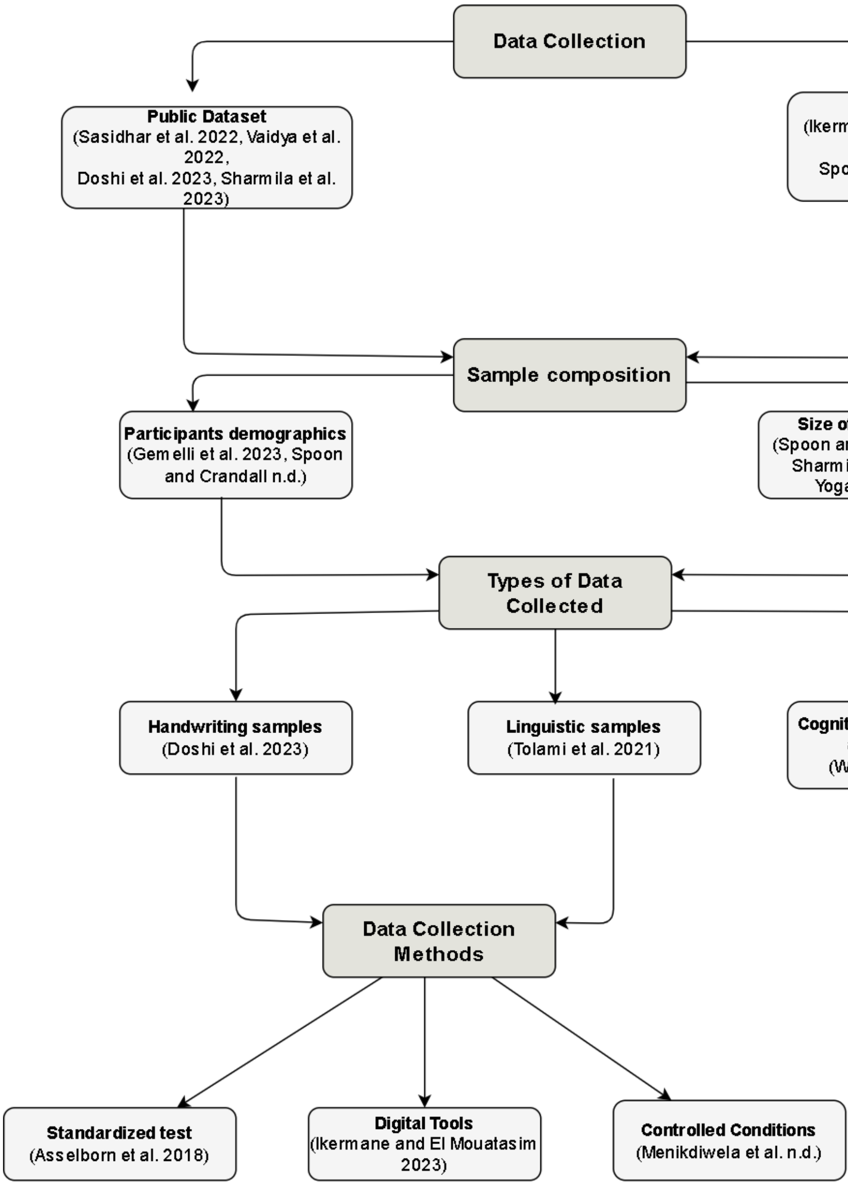


Figure 2-Data Acquisition Methodology

Ethical Considerations

Ethical considerations are crucial in data collection, particularly when handling sensitive health data. The two studies (Masood et al., 2023) and (Asselborn et al., 2018) have demonstrated robust ethical compliance by securing informed consent and obtaining ethical approval, underscoring their commitment to participant confidentiality and adherence to the Helsinki Declaration. In contrast, (Kunhoth et al., 2023b) and another study on CNN feature and classifier fusion by (Kunhoth et al., 2023a) mention only general compliance with ethical standards or use of publicly available data, suggesting reliance on prior ethical clearances. Another study (Walda et al., 2022) reported that no ethical review was required per local legislation, yet they ensured that written informed consent was obtained from the participants' legal guardians.

While only a small fraction of reviewed papers (5 out of 34) explicitly discuss ethical protocols, the variability in approaches reflects different requirements based on the study's nature and data sources. However, the importance of ethical considerations becomes particularly pronounced when dealing with sensitive data in the health sector. Maintaining strict ethical standards, especially when involving vulnerable populations like children, is essential for ensuring data integrity and protecting participants (Spoon & Crandall, n.d.).

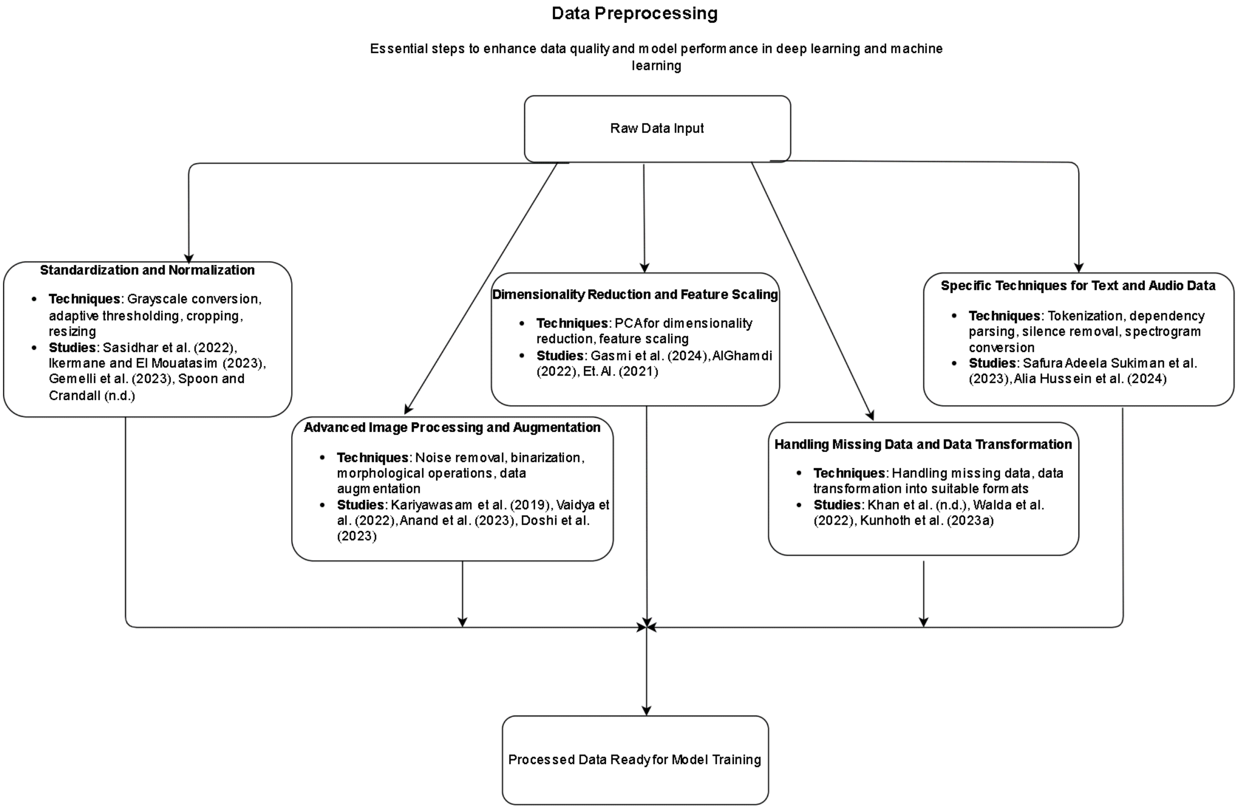
Data Preprocessing

Data preprocessing is crucial for preparing raw data for machine learning and deep learning models to improve data quality and model performance. Techniques like standardization and normalization are fundamental for consistency and suitability. For example, Sasidhar et al. (2022) used grayscale conversion and resizing on handwriting images, while Ikermane and El Mouatasim (2023) focused on normalizing features for handwriting analysis. Gemelli et al. (2023) and Spoon and Crandall (n.d.) also employed methods like digitization and cropping for standardization. Advanced image processing, such as noise removal and morphological operations, further enhances data utility. Kariyawasam et al. (2019) applied noise removal and normalization to both audio and image data. Data augmentation, used by Vaidya et al. (2022) and Anand et al. (2023), helps improve datasets' robustness. Doshi et al. (2023) utilized contrast enhancement and data augmentation to boost text recognition accuracy.

Dimensionality reduction, like Principal Component Analysis (PCA), helps reduce feature space and boost computational efficiency, as shown by Gasmi et al. (2024) and AlGhamdi (2022). As noted by Et, techniques like feature scaling ensure equal contribution of all features during model training. Al. (2021). Handling missing data is vital for dataset integrity, with strategies such as manual labelling and percentile calculations used by Khan et al. (n.d.). Walda et al. (2022) focused on transforming output variables for machine learning models. Data transformation for specific formats suitable for deep learning is also critical, as Kunhoth et al. (2023a) demonstrated. For text and audio data, specific preprocessing steps are essential. Safura Adeela Sukiman et al. (2023) employed techniques like tokenization for text data cleaning, while Alia Hussein et al. (2024) used silence removal and spectrogram conversion to prepare audio data for CNN models.

Fig. 3 shows the Comprehensive Overview of Data Preprocessing Techniques for ML and DL Models Based on Filtered Articles.

Figure 3-Comprehensive Overview of Data Preprocessing Techniques for Machine Learning and Deep Learning Models, Based on Filtered Articles



Feature Extraction and Selection

Feature extraction and selection are crucial in predicting Dysgraphia through handwriting recognition, enhancing the classification accuracy by refining raw data into pertinent features. Advanced models like convolutional neural networks (CNNs) are extensively used because they can extract spatial features from handwriting images. For example, Sasidhar et al. (2022) utilized a ResNet-50 model for high-level feature extraction in dyslexia detection, while Ikermane and El Mouatasim (2023) captured static, kinematic, pressure, and tilt features using a Wacom Intuos tablet. Further, (KT et al., 2023) combined CNNs with LSTM to capture both spatial and temporal features, and Liu et al. (2024) used CNNs with positional encoding and Bi-LSTM for comprehensive spatial, sequential, and positional feature integration, enhancing the detection capabilities for Dyslexia (Dara & Tumma, 2018). Additionally, some studies specifically tailored their feature extraction techniques to their tasks. Kariyawasam et al. (2019) implemented a voice recognition CNN to extract MFCC features from audio data and a separate CNN for handwriting analysis, aiding in comprehensive screening and intervention for various learning disabilities.

Feature selection optimizes these extracted features to improve model efficiency and performance. Techniques such as PCA help reduce the feature set effectively, as demonstrated by Gasmi et al. (2024) and AlGhamdi (2022), who streamlined features for better dyslexia prediction. Moreover, deep learning networks often incorporate feature selection mechanisms within their architecture; for instance, Gemelli et al. (2023) employed ResNet18 to merge traditional and deep learning features robustly. Specific algorithms like Random Forest and Elastic Net are also used to enhance selection accuracy, as seen in studies by Asselborn et al. (2018) and Et. Al. (2021).

Classification and Performance of Deep Learning Models

Fig. 4 presents the different classification models utilized for detecting Dyslexia and Dysgraphia, highlighting the diverse approaches in this field. Fig 5. expands on this by detailing the subdivisions of each classification model, illustrating the specific methodologies employed within the given framework.

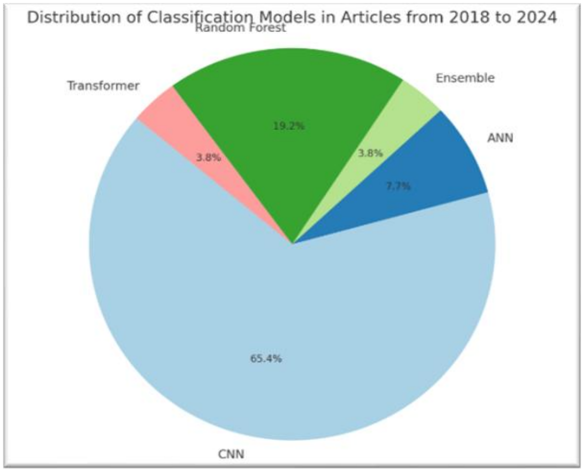


Figure 4-Distribution of Classification Models in Articles Published Between 2018 and 2024

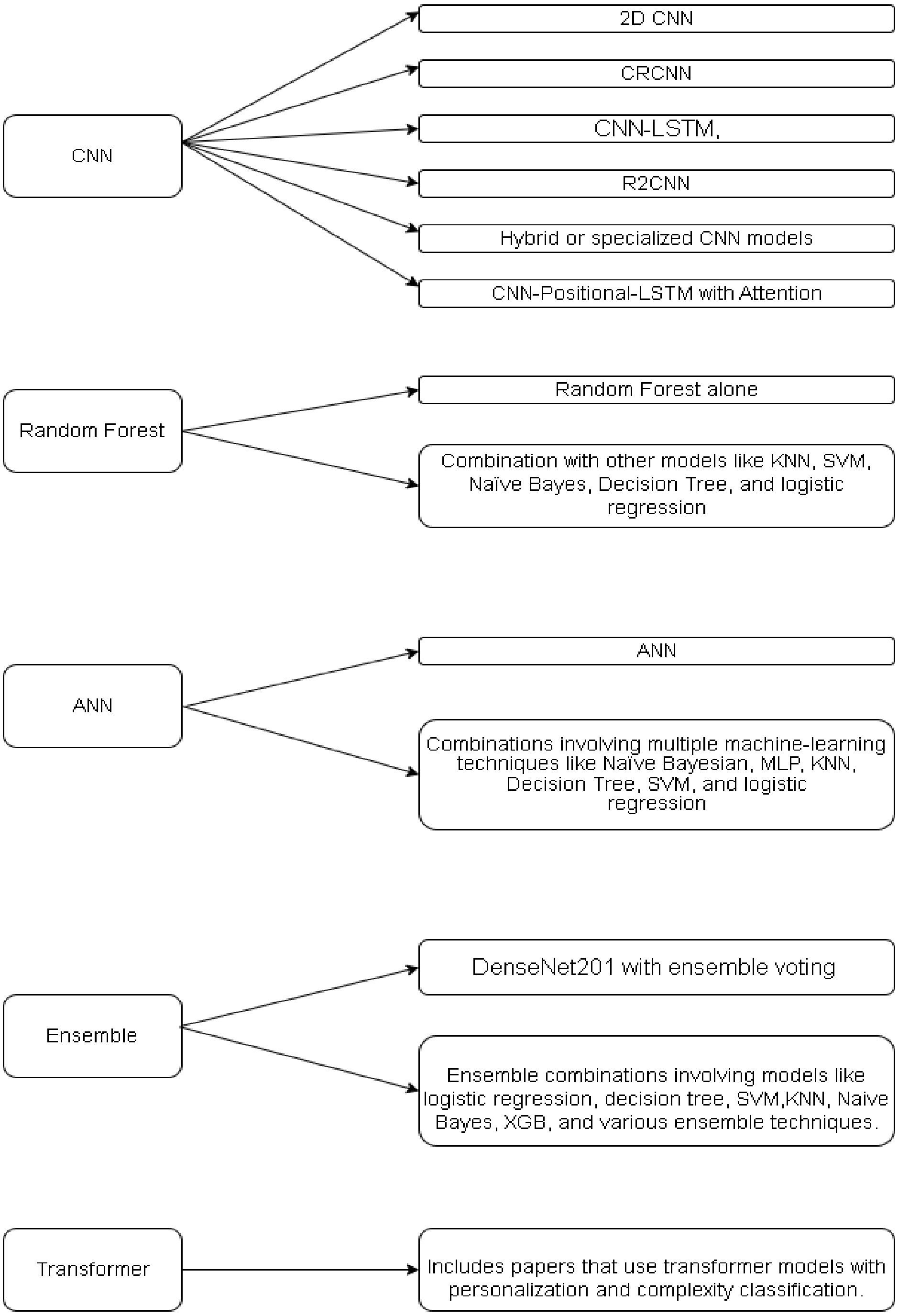


Figure 5-Subdivisions of each Classification Model

Performance Evaluation

The CNN-SVM model, as reported by Alqahtani et al. (2023), exemplifies the highest accuracy of 99.33%, outperforming other models by combining the robust feature extraction capabilities of CNN with the precise classification power of SVM. This hybrid approach effectively addresses challenges such as small data sizes and class imbalances, which have previously limited the accuracy of DL applications in dyslexia detection. Other ensemble methods like Random Forest, KNN, and Decision Tree also demonstrate high effectiveness, achieving up to 99.33% accuracy. This paper underscores the potential of combining DL and ML techniques to enhance the predictive accuracy and reliability of dyslexia diagnostics through handwriting analysis. Specific models for conditions like numeric Dysgraphia (Asselborn et al., 2018) showed an accuracy of 99.2%. Moderate accuracies were observed in models like AdaBoost (Kunhoth et al., 2023b) (AlGhamdi, 2022), ranging from 80.8% to 90.3%, suggesting careful tuning. Other methods, including validation accuracy metrics and soft voting ensembles, showed a reliable performance of around 90.4%. These findings affirm the potential of AI-based methods to enhance detecting and managing Dyslexia and related learning difficulties significantly.

In our analysis of the "Pubudu" screening tool's performance (Kariyawasam et al., 2019), dyslexia detection exhibited a notably lower accuracy of 65% when applied to a cohort of 50 students, 20 of whom were pre-diagnosed with Dyslexia. This contrasts with the higher accuracies observed in other screening processes, such as Letter dysgraphia (85%), Dyscalculia (90%), and Numeric Dysgraphia (90%). The diminished effectiveness in dyslexia screening could largely be attributed to linguistic variations; specifically, the tool achieved a lower precision of 58% for Sinhala, likely due to limited training data and smaller dataset sizes. In contrast, the English language screening demonstrated a significantly higher accuracy of 94%, benefiting from a more robust dataset. This variance underscores the critical impact of dataset comprehensiveness on the diagnostic accuracy of language-dependent tools like "Pubudu"(Kariyawasam et al., 2019). Fig. 6 shows the accuracy of the different papers from 2018-2024.

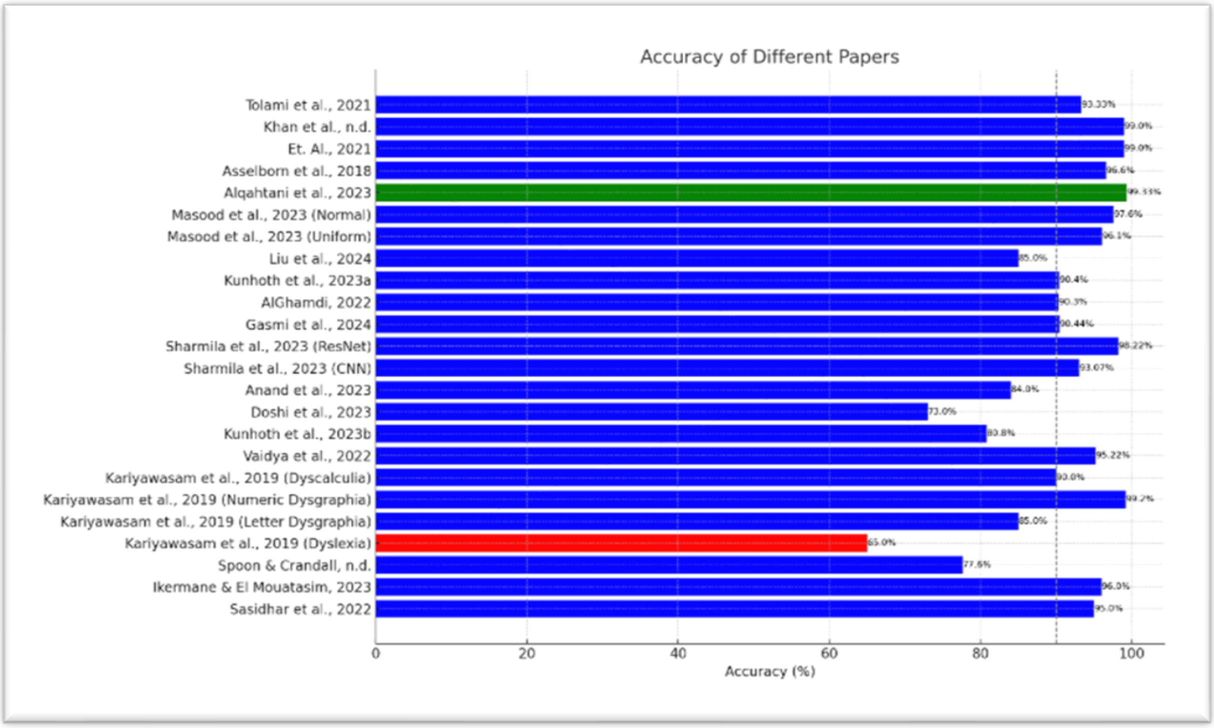


Figure 6-Accuracy of Predictive Models in Papers Published between 2018 and 2024

Discussion

While ML and DL are prominent techniques for diagnosing Dysgraphia and Dyslexia, research indicates the existence of various other methods for detection. These alternative approaches provide essential insights into the diagnostic tools and cognitive characteristics of Dysgraphia, which are fundamental for devising effective treatments. This variety raises an important question: Are machine learning and deep learning the most effective methods for detecting these learning disabilities? In the

Research Rabbit analysis, we discovered additional research papers focused on detecting Dysgraphia using traditional methods. These studies do not incorporate artificial intelligence approaches.

For instance, the development of the TestGraphia software automates dysgraphia diagnosis through digital handwriting analysis, although its efficacy might be limited by the algorithm's ability to interpret diverse handwriting nuances (Dimauro et al., 2020). Similarly, an expert system employing forward chaining has been introduced, simplifying the diagnosis process but relying heavily on its rule base's comprehensiveness, potentially missing atypical dysgraphia cases (Kurniawan et al., 2017). Another research (McCloskey & Rapp, 2017) offers a cognitive framework for understanding developmental Dysgraphia grounded in neuropsychology, yet it requires further empirical testing for broader educational applicability. Applying the Tunable Q-factor Wavelet Transform introduces a novel signal processing technique to identify dystrophic handwriting, needing validation across larger datasets to confirm its diagnostic reliability (Zvoncak et al., 2019). These studies collectively advance our understanding of Dysgraphia, highlighting the need for integrating various methods to enhance diagnostic accuracy and treatment efficacy.

When examining research on Dysgraphia, it is clear that two primary approaches dominate the field:

- The traditional methods concentrate on cognitive and neuropsychological evaluations.
- Contemporary strategies utilize ML and deep DL techniques.

Traditional methods are grounded in psychological and cognitive theories, providing detailed, individualized assessments considering various cognitive, psychological, and educational factors. However, these methods are time-consuming, subjective, and difficult to scale. In contrast, ML and DL approaches leverage algorithms to analyze large datasets quickly and objectively, identifying subtle handwriting patterns indicative of Dysgraphia. These techniques are efficient and suitable for large-scale screenings but often lack a theoretical foundation and can struggle with data dependency and generalization issues. The primary distinction between these approaches is that their traditional methods emphasize theoretical understanding, while ML/DL focuses on data-driven insights.

Combining these approaches could offer a more comprehensive method for diagnosing and understanding Dysgraphia, improving both the efficiency of large-scale screenings and the depth of individual assessments. Such integration could lead to more effective diagnostic and therapeutic strategies, utilizing the strengths of both traditional insights and technological advancements.

Recommendations

Research on Dyslexia and Dysgraphia predominantly centres on Latin-based languages, overlooking those with non-Latin scripts like Sinhala, Tamil, Hindi, and Chinese. This gap hinders understanding these conditions' unique challenges and markers in non-Western contexts. Therefore, there is a need to focus on English and other languages, developing models that accurately identify linguistic features. Enhancing linguistic diversity in training datasets is crucial, especially for learning disabilities linked to handwriting, to improve global applicability and diagnostic accuracy.

Beyond handwriting, Dysgraphia involves various factors affecting children with the condition. One of the studies (Kunhoth et al., 2024) expanded dysgraphia analysis to include cognitive abilities, motor skills, and language and phonological skills. This comprehensive approach aids in a deeper understanding and diagnosis of Dysgraphia:

- Cognitive abilities: Assessments in children focus on retrieval fluency, intelligence, working memory, and executive functions, which are critical for cognitive processing and writing tasks (Chung et al., 2020).
- Motor skills: The Beery Visual Motor Integration Test assesses visual-motor integration through tasks requiring students to replicate complex images, highlighting the link between motor coordination and writing (Sutton et al., 2011).
- Language and phonological skills: Crucial for dysgraphia diagnosis, these skills are evaluated through various tests, assessing phonological awareness and handwriting quality, essential for spelling and writing (Barnett et al., 2009; Korkman et al., 2012; Wagner et al., 2016).

The study (Ikermane & El Mouatasim, 2023) noted that boys are more likely than girls to have Dysgraphia, suggesting that factors beyond cognitive, motor, and phonological skills can influence the dysgraphic condition. This observation underscores the importance of considering diverse personal information (multi-model) in dysgraphia analysis. In this context, a multi-model approach is highly beneficial as it combines various data sources, like text and handwriting images, to improve machine learning tasks in detecting Dysgraphia through handwriting. By integrating these multiple modalities, future research could offer a more comprehensive understanding and detection of Dysgraphia, thereby expanding the scope and depth of scholarly inquiries into this field.

Limitations

CNN models frequently face overfitting during training, where they perform exceptionally well on training sets but poorly on unseen testing sets due to complex, parameter-heavy architectures (Khan et al., 2020). The limited size of many Dyslexia datasets often worsens this problem. To mitigate overfitting, various strategies are employed, such as:

- **Data augmentation** effectively increases the dataset size using techniques such as cropping, translation, rotation, and noise injection, which also help balance class distribution (Sasidhar et al., 2022; Vaidya et al., 2022; Sharmila et al., 2023; Doshi et al., 2023).
- **Dropout** is another technique used during training that randomly deactivates neurons, promoting the learning of more robust and generalized features by spreading the feature selection process across a more comprehensive array of neurons (Doshi et al., 2023; Anand et al., 2023).
- **Cross-validation and proper train/test splits** are essential for verifying the model's performance on independent datasets, ensuring effective generalization to new, unseen data (Yogarajah, n.d.).

Conclusion

This review summarized DL and ML-based dysgraphia detection techniques and examined the critical factors for building accurate predictive models. Selected articles for this analysis were identified through major academic databases using specific search keywords and criteria, following the PRISMA flow diagram and pre-defined inclusion criteria. Ultimately, 34 machine learning and deep learning studies focused on dysgraphia detection were thoroughly reviewed. The research shows that leveraging both public and custom datasets enables a comprehensive understanding of Dysgraphia, providing robust training data for developing predictive models. Additionally, the review emphasizes the importance of adhering to stringent ethical standards, especially when handling sensitive data, to ensure participant confidentiality and maintain the integrity of the research process. Essential preprocessing steps like standardization, normalization, advanced image processing, and data augmentation enhance data quality and model performance. Feature extraction and selection, mainly using CNNs and dimensionality reduction techniques like PCA, are critical for capturing relevant attributes and ensuring efficient model training. Future research should emphasize the collection of dysgraphia-related datasets to enhance model accuracy and efficacy.

Future Directions

Future research should focus on several key areas to enhance the accuracy and applicability of dysgraphia prediction models:

***Multimodal Deep Learning Architectures**:*

Explore architectures that integrate handwriting data with cognitive details, demographic information, and other relevant data for a comprehensive and accurate prediction of Dysgraphia.

Linguistic Diversity:

Extend research beyond English to include other languages, developing models that account for linguistic and cultural differences in handwriting and cognitive processing for global applicability.

Funding: The research conducted did not receive any external funding.

Availability of Data: This review paper does not require collecting new data; instead, it assesses dyslexia-related data from 34 papers. The dataset on dyslexia-related handwriting can be accessed at https://www.kaggle.com/datasets/drizasazanitaisa/dyslexia-handwriting-dataset. Other research that utilized custom handwriting datasets is not publicly available but can be requested by the authors.

Conflict of Interest: The authors confirm that no conflict of interest is involved.

References

AlGhamdi, A. S. (2022). Novel Ensemble Model Recommendation Approach for the Detection of Dyslexia. *Children*, *9*(9), 1337. https://doi.org/10.3390/children9091337

Alia Hussein, Ahmed Talib Abdulameer, Ali Abdulkarim, Husniza Husni, & Dalia Al-Ubaidi. (2024). Classification of Dyslexia Among School Students Using Deep Learning. *Journal of Techniques*, *6*(1), 85–92. https://doi.org/10.51173/jt.v6i1.1893

Alqahtani, N. D., Alzahrani, B., & Ramzan, M. S. (2023). Detection of Dyslexia Through Images of Handwriting using Hybrid AI Approach. *International Journal of Advanced Computer Science and Applications*, *14*(10). https://doi.org/10.14569/IJACSA.2023.0141099

Anand, S., Sharma, P., Sharma, M., Talwar, R., & Mongia, C. (2023). *Deep Learning-Based Classification of Dysgraphia Severity in Children from Handwriting Images*. https://doi.org/10.21203/rs.3.rs-3803290/v1

Asselborn, T., Gargot, T., Kidziński, Ł., Johal, W., Cohen, D., Jolly, C., & Dillenbourg, P. (2018). Automated human-level diagnosis of Dysgraphia using a consumer tablet. *Npj Digital Medicine*, *1*(1), 42. https://doi.org/10.1038/s41746-018-0049-x

Barnett, A. L., Henderson, S. E., Scheib, B., & Schulz, J. (2009). Development and standardization of a new handwriting speed test: The Detailed Assessment of Speed of Handwriting. In V. Connelly, A. L. Barnett, J. E. Dockrell, & A. Tolmie (Eds.), *BJEP Monograph Series II: Part 6 Teaching and Learning Writing*. British Psychological Society. https://doi.org/10.1348/000709909X421937

Chung, P. J., Patel, D. R., & Nizami, I. (2020). Disorder of written expression and Dysgraphia: Definition, diagnosis, and management. *Translational Pediatrics*, *9*(S1), S46–S54. https://doi.org/10.21037/tp.2019.11.01

Dara, S., & Tumma, P. (2018). Feature Extraction By Using Deep Learning: A Survey. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 1795–1801. https://doi.org/10.1109/ICECA.2018.8474912

Dimauro, G., Bevilacqua, V., Colizzi, L., & Di Pierro, D. (2020). TestGraphia, a Software System for the Early Diagnosis of Dysgraphia. *IEEE Access*, *8*, 19564–19575. https://doi.org/10.1109/ACCESS.2020.2968367

Doshi, N. N., Maniyar, M. U., Shah, K. K., Sarda, N. D., Narvekar, M., & Mukhopadhyay, D. (2023). A Convolutional Recurrent Neural Network-Based Model For Handwritten Text Recognition To Predict Dysgraphia. *2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS)*, 145–150. https://doi.org/10.1109/ICISCoIS56541.2023.10100514

Dysgraphia Classification based on the Non-Discrimination Regularization in Rotational Region Convolutional Neural Network. (2022). *International Journal of Intelligent Engineering and Systems*, *15*(1). https://doi.org/10.22266/ijies2022.0228.06

Et. Al., S. D. (2021). Comparison of Classification Methods Used in Machine Learning for Dysgraphia Identification. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(11), 1886–1891. https://doi.org/10.17762/turcomat.v12i11.6142

Gasmi, K., Ammar, L. B., Krichen, M., Alamro, M. A., Mihoub, A., & Mrabet, M. (2024). Optimal Ensemble Learning model for Dyslexia prediction based on an adaptive genetic algorithm. *IEEE Access*, 1–1. https://doi.org/10.1109/ACCESS.2024.3395803

Gemelli, A., Marinai, S., Vivoli, E., & Zappaterra, T. (2023). Deep-learning for dysgraphia detection in children handwritings. *Proceedings of the ACM Symposium on Document Engineering 2023*, 1–4. https://doi.org/10.1145/3573128.3609351

Hewapathirana, C., Abeysinghe, K., Maheshani, P., Liyanage, P., Krishara, J., & Thelijjagoda, S. (2021). A Mobile-Based Screening and Refinement System to Identify the Risk of Dyscalculia and Dysgraphia Learning Disabilities in Primary School Students. *2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*, pp. 287–292. https://doi.org/10.1109/ICIAfS52090.2021.9605998

Ikermane, M., & El Mouatasim, A. (2023). Digital handwriting characteristics for dysgraphia detection using artificial neural network. *Bulletin of Electrical Engineering and Informatics*, *12*(3), 1693–1699. https://doi.org/10.11591/eei.v12i3.4571

Kariyawasam, R., Nadeeshani, M., Hamid, T., Subasinghe, I., Samarasinghe, P., & Ratnayake, P. (2019). Pubudu: Deep Learning Based Screening And Intervention of Dyslexia, Dysgraphia And Dyscalculia. *2019 14th Conference on Industrial and Information Systems (ICIIS)*, 476–481. https://doi.org/10.1109/ICIIS47346.2019.9063301

Khan, A., Sohail, A., Zahoora, U., & Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, *53*(8), 5455–5516. https://doi.org/10.1007/s10462-020-09825-6

Korkman, M., Kirk, U., & Kemp, S. (2012). *NEPSY - Second Edition* [Dataset]. https://doi.org/10.1037/t15125-000

K.T., J., V., L., & P., J. A. (2023). DyslexiScan: A Dyslexia Detection Method from Handwriting Using CNN LSTM Model. *2023 International Conference on Innovations in Engineering and Technology (ICIET)*, 1–6. https://doi.org/10.1109/ICIET57285.2023.10220750

Kunhoth, J., Al Maadeed, S., Saleh, M., & Akbari, Y. (2023a). CNN feature and classifier fusion on novel transformed image dataset for dysgraphia diagnosis in children. *Expert Systems with Applications*, *231*, 120740. https://doi.org/10.1016/j.eswa.2023.120740

Kunhoth, J., Al Maadeed, S., Saleh, M., & Akbari, Y. (2023b). Exploration and analysis of On-Surface and In-Air handwriting attributes to improve dysgraphia disorder diagnosis in children based on machine learning methods. *Biomedical Signal Processing and Control*, *83*, 104715. https://doi.org/10.1016/j.bspc.2023.104715

Kunhoth, J., Al-Maadeed, S., Kunhoth, S., Akbari, Y., & Saleh, M. (2024). Automated systems for diagnosis of Dysgraphia in children: A survey and novel framework. *International Journal on Document Analysis and Recognition (IJDAR)*. https://doi.org/10.1007/s10032-024-00464-z

Kurniawan, D. A., Sihwi, S. W., & Gunarhadi. (2017). An expert system for diagnosing Dysgraphia. *2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 468–472. https://doi.org/10.1109/ICITISEE.2017.8285552

Liu, H. W., Wang, S., & Tong, S. X. (2024). DysDiTect: Dyslexia Identification Using CNN-Positional-LSTM-Attention Modeling with Chinese Dictation Task. *Brain Sciences*, *14*(5), 444. https://doi.org/10.3390/brainsci14050444

Lopez, C., Hemimou, C., Golse, B., & Vaivre-Douret, L. (2018). Developmental Dysgraphia is often associated with minor neurological dysfunction in children with developmental coordination disorder (DCD). *Neurophysiologie Clinique*, *48*(4), 207–217. https://doi.org/10.1016/j.neucli.2018.01.002

Masood, F., Khan, W. U., Ullah, K., Khan, A., Alghamedy, F. H., & Aljuaid, H. (2023). A Hybrid CNN-LSTM Random Forest Model for Dysgraphia Classification from Handwritten Characters with Uniform/Normal Distribution. *Applied Sciences*, *13*(7), 4275. https://doi.org/10.3390/app13074275

McCloskey, M., & Rapp, B. (2017). Developmental Dysgraphia: An overview and framework for research. *Cognitive Neuropsychology*, *34*(3–4), pp. 65–82. https://doi.org/10.1080/02643294.2017.1369016

Menikdiwela, K., Bulathwatte, A., Vojtova, V., & Pathirana, B. (n.d.). *Facilitating individuals with specific learning disorders in Sri Lanka and the Czech Republic*.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Journal of Clinical Epidemiology*, *62*(10), 1006–1012. https://doi.org/10.1016/j.jclinepi.2009.06.005

Patil, S., Apare, R., Borhade, R., & Mahalle, P. (2024). Intelligent approaches for early prediction of learning disabilities in children using learning patterns: A survey and discussion. *Journal of Autonomous Intelligence*, *7*(5), 1329. https://doi.org/10.32629/jai.v7i5.1329

Safura Adeela Sukiman, Nor Azura Husin, Hazlina Hamdan, & Masrah Azrifah Azmi Murad. (2023). A Hybrid Personalized Text Simplification Framework Leveraging the Deep Learning-based Transformer Model for Dyslexic Students. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *34*(1), 299–313. https://doi.org/10.37934/araset.34.1.299313

Sasidhar, A., Kumar, G. K., Yoshitha, K., & Tulasi, N. (2022). Dyslexia Discernment in children based on Handwriting images using Residual Neural Network. *2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, 1–6. https://doi.org/10.1109/CSITSS57437.2022.10026368

Sharmila, C., Shanthi, N., Santhiya, S., Saran, E., Sri Rakesh, K., & Sruthi, R. (2023). An Automated System for the Early Detection of Dysgraphia Using Deep Learning Algorithms. *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, pp. 251–257. https://doi.org/10.1109/ICSCDS56580.2023.10105022

Spoon, K., & Crandall, D. (n.d.). *Can We (and Should We) Use AI to Detect Dyslexia in Children's Handwriting?*

Sutton, G. P., Barchard, K. A., Bello, D. T., Thaler, N. S., Ringdahl, E., Mayfield, J., & Allen, D. N. (2011). Beery-Buktenica Developmental Test of Visual-Motor Integration performance in children with traumatic brain injury and attention-deficit/hyperactivity disorder. *Psychological Assessment*, *23*(3), 805–809. https://doi.org/10.1037/a0023370

Tao, Y., & Rapp, B. (2019). The effects of lesion and treatment-related recovery on functional network modularity in post-stroke Dysgraphia. *NeuroImage: Clinical*, *23*, 101865. https://doi.org/10.1016/j.nicl.2019.101865

Tolami, F. A., Khorasani, M., Kahani, M., Yazdi, S. A. A., & Ghalenoei, M. A. (2021). An intelligent linguistic error detection approach to automated diagnosis of Dyslexia disorder in Persian-speaking children. *2021 11th International Conference on Computer Engineering and Knowledge (ICCKE)*, 393–398. https://doi.org/10.1109/ICCKE54056.2021.9721446

Vaidya, V., Pravanth, T., & Viji, D. (2022). Air Writing Recognition Application for Dyslexic People. *2022 International Mobile and Embedded Technology Conference (MECON)*, 553–558. https://doi.org/10.1109/MECON53876.2022.9752119

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2016). *Comprehensive Test of Phonological Processing–Second Edition* [Dataset]. https://doi.org/10.1037/t52630-000

Walda, S., Hasselman, F., & Bosman, A. (2022). Identifying Determinants of Dyslexia: An Ultimate Attempt Using Machine Learning. *Frontiers in Psychology*, *13*, 869352. https://doi.org/10.3389/fpsyg.2022.869352

Yogarajah, P. (n.d.). *Deep Learning Approach to Automated Detection of Dyslexia-Dysgraphia*.

Zvoncak, V., Mekyska, J., Safarova, K., Smekal, Z., & Brezany, P. (2019). New Approach of Dysgraphic Handwriting Analysis Based on the Tunable Q-Factor Wavelet Transform. *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 289–294. https://doi.org/10.23919/MIPRO.2019.8756872