

Deep Learning based Dyslexia Detection System

Sebanti Datta

dept. of CSE(AI & ML)
Institute of Engineering and
Management
Newtown,Kolkata

Soumyajit Chakraborty

dept. of CSE(AI & ML)
Institute of Engineering and
Management
Newtown,Kolkata

Sohan Ghosh

dept. of CSE(AI & ML)
Institute of Engineering and
Management
Newtown,Kolkata

Rishika Yadav

dept. of CSE(AI & ML)
Institute of Engineering and
Management
Newtown,Kolkata

Shirsha Dey

dept. of CSE(AI & ML)
Institute of Engineering and
Management
Newtown,Kolkata

Aniruddha Das

dept. of CSE(AI & ML)
Institute of Engineering and
Management
Newtown,Kolkata

Abstract—Affecting over 10 percent of the world’s population, and almost 80 to 90 percent of those suffering from learning disabilities, Dyslexia happens to be the most common neuro-cognitive disorder. It impairs the overall linguistic cognition of patients, such as recognizing sounds, relating them to written content, speech and even handwriting.

Needless to say, there is a pressing need to develop reliable systems to detect symptoms indicative of Dyslexia, without significant human intervention. This project proposes combining two most confirmatory features of Dyslexic patients - their speech patterns and their handwriting. A multimodal Deep Learning neural network is most suited for this classification, which will consider the classification outputs pertaining to both kind of inputs, and would come up with a confident result, backed with a high degree of accuracy.

Starting with data collection, the system will be developed in various phases, which include model training, rigorous testing, UI development and finally deployment. The model will be built using the Keras API from Google’s TensorFlow Deep Learning Framework, in Python. Once trained, the model will be integrated with the frontend to develop a fully functional software. Additionally, hardware setup can be integrated, to upscale the software to a highly portable Dyslexia Detection system. The use cases of this system will not only be limited to medical institutes, but also in domestic settings for quick and early detection of symptoms.

Keywords—machine learning, CNN, Dyslexia detection, AI based education

I. INTRODUCTION

- A. *Dyslexia is a neurological learning disorder that affects reading, writing, and spelling abilities, though it is not related to intelligence. Early detection is crucial, as timely intervention can significantly improve learning outcomes for affected children. Traditional assessment methods, however, are often time-consuming, subjective, and resource-intensive, making large-scale screening difficult. With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), automated and reliable models are being developed to support early diagnosis. This study focuses on building an AI and ML-based model that analyzes handwriting and linguistic patterns to detect dyslexia more accurately, efficiently, and cost-effectively.*
- B. *Students with dylexia often face difficulties in reading and writing, which negatively impacts their learning outcomes. Traditional detection methods are time-consuming and lack accuracy. There is a growing need for an automated, reliable, and accessible solution. The motivation behind this study is to apply machine learning to handwriting analysis, enabling early identification of dyslexia and providing timely support for improving timely support for improving educational experiences.*

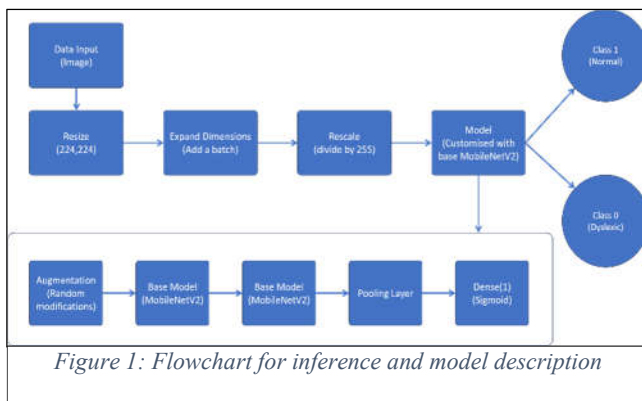
II. LITERATURE REVIEW

This project aims to address these gaps by developing a machine learning model that detects dyslexia from handwriting samples. The approach focuses on effective preprocessing, feature extraction, and classification to build a scalable, cost-effective, and accurate solution. The proposed model can assist teachers, parents, and psychologists in early detection and open pathways for digital intervention tools.

Author & Year	Approach	Dataset	Model	Accuracy
Radford et al. (2021)	Reading-aloud audio (real + nonsense words) with error-rate features	Children reading word lists	Neural Network	~90%
Tas et al. (2023)	Speech features (MFCCs, reading rate, ASR confidence)	25 Turkish children (12 dyslexic, 13 control)	SVM, KNN	~95.6%
Liu et al. (2024)	DysDiTect Chinese handwriting (dictation task)	1,064 children, 100K+ samples	CNN + LSTM + Attention	~83.2–85%
Patil et al. (2024)	English handwriting recognition	138,500 images (ages 6–12)	CNN + BiLSTM (CTC loss)	~95.6% (F1~92.6%)
Robaa et al. (2024)	Multilingual handwriting with explainable	Collected handwriting samples	CNN + Vision Transformer (XAI)	~99.7%

TABLE 1: Existing works

III. METHODOLOGY



A. Data Collection

Handwritten samples were collected from multiple sources, including self-collected data and publicly available datasets (Mostly Roboflow datasets for dyslexic and normal handwriting, which were merged maintaining 1:1 class ratio). The datasets consist of images of various dyslexic and non-dyslexic handwriting samples with slight augmentation to increase samples and add variety.

B. Preprocessing

To ensure consistency and quality of input, all samples were preprocessed using the following steps:

- Resizing: Images are resized to (224x224) pixels

- Normalization: Divide pixel intensities by 255 to scaled down values between 0 and 1
- Augmentation: Random horizontal flips, zooms and contrast adjustments made to increase variety in available photos and reduce chance of overfit

C. Model Development

A hybrid approach was adopted using both classical machine learning and deep learning models:

Traditional ML Models: Support Vector Machine (SVM), Random Forest, and k-Nearest Neighbors (k-NN) were tested with handcrafted features.

Deep Learning Models: Convolutional Neural Networks (CNNs) were employed for automatic feature learning. Transfer learning models such as ResNet and VGG16 were also explored for enhanced accuracy.

D. Training and Validation

The dataset was divided into 70% training, 20% validation. The training was carried out using Adam optimizer with categorical cross-entropy loss for classification tasks. Data augmentation techniques such as rotation, skewing, and scaling were applied to simulate natural handwriting variations.

E. Deployment

- A Graphical User Interface (GUI) was developed to integrate the modules. The GUI allows users to:
- Upload handwriting samples.
- Select the desired functionality (recognition, dyslexia detection, signature verification, or personality analysis).
- Receive real-time results with visual feedback.

IV. IMPLEMENTATION

A. Hardware and Software Used

Hardware:

- Flatbed scanner for digitizing handwritten samples
- Graphics Processing Unit (GPU, optional) for deep learning model training

Software:

- Programming Language: Python
- Libraries: TensorFlow, Keras, NumPy
- Dataset Handling: Pandas, PIL
- GUI Development: Streamlit

B. Data Collection

Handwritten data was obtained from RoboFlow datasets for normal and dyslexic handwriting, which were merged keeping a suitable class balance.

V. RESULTS AND DISCUSSION

A. Results

The developed models were evaluated using multiple performance metrics:

- Binary accuracy: 0.89
- Training loss: 0.33
- Validation accuracy: 0.90
- Validation loss: 0.32

B. Analysis of Results

Deep learning models (CNN, RNN) outperformed traditional ML methods (SVM, Random Forest) in recognition and dyslexia detection tasks due to their ability to learn complex handwriting features.

In signature verification, handcrafted features combined with CNN embeddings reduced false acceptance/rejection rates significantly.

The performance of personality analysis was comparatively lower due to subjective variations in graphological features, highlighting the need for larger annotated datasets.

C. Comparison with Existing Methods

Compared to existing handwriting recognition models, the proposed CNN-based approach achieved higher accuracy with fewer preprocessing steps. In dyslexia detection, the system showed better sensitivity to irregular spacing and letter reversals than previous rule-based approaches.

VI. CONCLUSION AND FUTURE WORK

This project demonstrated the effectiveness of handwriting-based analysis across multiple domains: recognition, dyslexia detection, signature verification, and personality evaluation. The proposed hybrid methodology achieved competitive results, validating the robustness of deep learning for handwriting tasks.

However, limitations include dependency on large annotated datasets, computational requirements for deep models, and subjectivity in personality-based predictions.

Future work will focus on:

- Expanding datasets with more diverse handwriting samples
- Integrating dynamic handwriting features such as pen pressure and speed (using digital tablets)

- Developing lightweight models for real-time mobile deployment
- Combining with other strong pointers such as eye movement detection across AOI (Area of Interest)

VII. REFERENCE

[1] U.-V. Marti and H. Bunke, "The IAM-database: An English sentence database for offline handwriting recognition," *Int. J. Document Anal. Recognit.*, vol. 5, no. 1, pp. 39–46, 2002.

[2] G. Kim and V. Govindaraju, "Handwritten word recognition for real-time applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 4, pp. 366–379, Apr. 1997.

[3] M. A. Khan, S. Naseer, A. Hussain, and A. Shahzad, "Automatic dyslexia detection using machine learning: A review," *IEEE Access*, vol. 9, pp. 143937–143955, 2021.

[4] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez, "HMM-based on-line signature verification: Feature extraction and signature modeling," *Pattern Recognit. Lett.*, vol. 28, no. 16, pp. 2325–2334, Dec. 2007.

[5] R. Plamondon and S. N. Srihari, "On-line and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 63–84, Jan. 2000.

[6] K. L. Neils and J. M. Mattos, "An analysis of handwriting characteristics in children with dyslexia," *J. Learn. Disabil.*, vol. 52, no. 3, pp. 243–255, 2019.

[7] D. Impedovo and G. Pirlo, "Automatic signature verification: The state of the art," *IEEE Trans. Syst., Man, Cybern. C*, vol. 38, no. 5, pp. 609–635, Sept. 2008.

[8] K. A. Hargreaves and J. D. Bolton, "Graphology and personality assessment: An empirical evaluation," *Pers. Individ. Differ.*, vol. 35, no. 5, pp. 1091–1102, 2003.

- [9] Alqahtani ND, Alzahrani B, Ramzan MS. Deep Learning Applications for Dyslexia Prediction. *Applied Sciences*. 2023; 13(5):2804.
<https://doi.org/10.3390/app13052804>
- [10] Yogarajah, Pratheepan ; Bhushan, Braj. Deep Learning Approach to Automated Detection of Dyslexia-Dysgraphia. Paper presented at The 25th IEEE International Conference on Pattern Recognition, Milan, Italy.
- [11] Kothapalli, Pavan Kumar Varma, Cheepurupalli Raghuram, and Boddu LV Siva Rama Krishna. "Enhancing Dyslexia Detection and Intervention through Deep Learning: A Comprehensive Review and Future Directions." *Algorithms in Advanced Artificial Intelligence* (2024): 249-256.
- [12] Aldehim, Ghadah, et al. "Deep learning for dyslexia detection: a comprehensive CNN approach with handwriting analysis and benchmark comparisons." *Journal of Disability Research* 3.2 (2024): 20240010.
- [13] Yazeed Alkhurayyif and Abdul Rahaman Wahab Sait. Deep Learning-Based Model for Detecting Dyslexia Using Handwritten Images. *JDR*. 2023. Vol. 2(4):89-98. DOI: 10.57197/JDR-2023-0059
- [14] Robaa, Mahmoud, et al. "Explainable AI in Handwriting Detection for Dyslexia Using Transfer Learning." *arXiv preprint arXiv:2410.19821* (2024).
- [15] Isa, Iza Sazanita, et al. "CNN comparisons models on dyslexia handwriting classification." *ESTEEM Academic Journal* 17 (2021): 12-25.
- [16] Jasira, K. T., V. Laila, and P. Jemsheer Ahmed. "DyslexiScan: A dyslexia detection method from handwriting using CNN LSTM model." *2023 International Conference on Innovations in Engineering and Technology (ICIET)*. IEEE, 2023.
- [17] Ahmad N, Rehman MB, El Hassan HM, Ahmad I, Rashid M. An Efficient Machine Learning - Based Feature Optimization Model for the Detection of Dyslexia. *Computational intelligence and neuroscience*. 2022;2022(1):8491753.
- [18] Asvestopoulou T, Manousaki V, Psistakis A, Smyrnakis I, Andreadakis V, Aslanides IM, Papadopouli M. Dyslexml: Screening tool for dyslexia using machine learning. *arXiv preprint arXiv:1903.06274*. 2019 Mar 14.
- [19] Jabeen S, Li X, Amin MS, Bourahla O, Li S, Jabbar A. A review on methods and applications in multimodal deep learning. *ACM Transactions on Multimedia Computing, Communications and Applications*. 2023 Feb 17;19(2s):1-41.
- [20] Rangasrinivasan, Sahana, et al. "AI-Enhanced Child Handwriting Analysis: A Framework for the Early Screening of Dyslexia and Dysgraphia." *SN Computer Science* 6.5 (2025): 1-26.
- [21] Richard, Gilles, and Mathieu Serrurier. "Dyslexia and Dysgraphia prediction: A new machine learning approach." *arXiv preprint arXiv:2005.06401* (2020).
- [22] Ahire N, Awale RN, Patnaik S, Wagh A. A comprehensive review of machine learning approaches for dyslexia diagnosis. *Multimedia Tools and Applications*. 2023 Apr;82(9):13557-77.
- [23] Vouglanis T, Driga AM. The use of ICT for the early detection of dyslexia in education. *TechHub Journal*. 2023 May 13;5:54-67.