Group05

Asim Mahmudov, Sima Maniyeva

Research Project

Deep Learning-Based Detection and Assessment of Dyslexia Using Machine Learning Models

Description of the Topic

Dyslexia is a neurologically based specific learning disorder that causes difficulties in reading, writing, and spelling skills in children. It is not related to IQ level, and children may have normal or high general intelligence, but the way the brain processes language and writing is different. When dyslexia is detected early and treated with the right methods, children can learn normally and successfully. Early diagnosis is crucial, yet traditional diagnostic methods—such as standardized reading tests, cognitive assessments, and clinical observations—are often time-consuming, resource-intensive, and subject to human bias. Moreover, these methods are not always accessible in under-resourced schools or remote regions, creating a gap in early intervention efforts.

To address these limitations, this study investigates the potential of deep learning models, particularly Convolutional Neural Networks (CNNs), in automatically detecting dyslexia using handwritten text images. Handwriting, due to its inherent irregularities in line structure, letter formation, and spacing, offers visual patterns that may serve as indicators of dyslexia. By analyzing these subtle variations, CNNs can learn to differentiate between dyslexic and non-dyslexic handwriting with minimal manual intervention.

The research utilizes a publicly available dataset containing handwritten English alphabet samples from both dyslexic and non-dyslexic individuals. A pre-trained VGG16 model, originally trained on the ImageNet dataset, is fine-tuned for this specific classification task. The model's top layers are modified, and additional dense layers are incorporated to adapt the network for binary classification. Transfer learning is employed to leverage the general visual understanding of the pre-trained model, while data augmentation techniques are used to artificially expand the dataset and improve generalization.

The primary objective of this research is to develop a robust and reliable deep learning pipeline capable of identifying visual markers of dyslexia with high accuracy. In doing so, we aim to automate the diagnostic process and reduce reliance on traditional methods. The model's performance is rigorously evaluated using various metrics such as accuracy, precision, recall, and F1-score to ensure its practical viability.

By validating and building upon the findings of Aldehim et al. (2024), this study not only contributes to the growing body of literature on AI-assisted dyslexia detection but also moves a step closer to real-world deployment. Ultimately, our long-term vision is to create an accessible, AI-driven early screening tool that can assist educators, clinicians, and families—especially in multilingual and resource-constrained environments—by providing timely and reliable identification of dyslexia.

Why This Topic?

- Dyslexia affects millions globally, often going undiagnosed until later school years. According to Aldehim et al. (2024), computer vision techniques — particularly convolutional neural networks — offer promising performance in identifying dyslexic handwriting characteristics such as letter deformation, irregular spacing, and inconsistent stroke flow.
- By applying a VGG16-based CNN, we aim to validate the findings of the referenced study in a reproducible and scalable pipeline, thereby extending its practical relevance.

National and Global Impact

- National Impact: Early diagnosis of dyslexia in schools can improve literacy rates and academic performance, leading to better educational outcomes. (According to us. We suppose it can impact with good way to improve academic performance) [1]
- Global Impact: About 10% of the world's population suffers from this disorder, and it is important to discover it early to reduce its impact and improve the skills of dyslexic. [3]
- An AI-based diagnostic tool can be deployed in various languages and regions, enabling cost-effective and widespread screening, particularly in under-resourced areas.

Scientific Impact

- By enabling accurate and early detection of dyslexia, the model facilitates timely interventions, which are crucial for improving academic and social outcomes for individuals with dyslexia. [4]
- Many recent studies and reviews have shown how different CNN model architectures have been used to analyze MRI datasets of various critical brain disorders such as brain tumor, Alzheimer's disease, Parkinson's disease and schizophrenia.[2]
- Improving the effectiveness of assistive technologies through AI-driven analysis.

Data Sources

1. Handwriting Dataset: This dataset contains handwriting data
   https://www.kaggle.com/datasets/oussamaslmani/dyslexic
   We use a VGG16 backbone pre-trained on ImageNet, followed by Global Average Pooling and fully connected dense layers tailored for binary classification.
   We found 3 dataset from kaggle but 2 of them are very large and we choose 1st dataset
2. Predicting Risk of Dyslexia - PLOS ONE
   https://www.kaggle.com/datasets/luzrello/dyslexia
3. Dyslexia Handwriting Dataset
   https://www.kaggle.com/datasets/drizasazanitaisa/dyslexia-handwriting-dataset

Methodology for Testing and Validation

- **Model Backbone**: VGG16 (ImageNet weights) without top layers
- **Fine-Tuning:** Added Global Average Pooling and Dense layers as per Aldehim et al.'s CNN structure
- **Data Augmentation:** Rotation, shifting, zoom, flipping using ImageDataGenerator
- **Evaluation Metrics:** Accuracy, Precision

Research Papers: References (We decide to use this research our main reference for our research work)

[1] Ghadah Aldehim, Mamoon Rashid and Ala Saleh Alluhaidan et al. Deep Learning for Dyslexia Detection: A Comprehensive CNN Approach with Handwriting Analysis and Benchmark Comparisons. *JDR.* 2024. Vol. 3(2). DOI: 10.57197/JDR-2024-0010

[2] OPEYEMI LATEEF USMAN, RAVIE CHANDREN MUNIYANDI, KHAIRUDDIN OMAR, AND MAZLYFARINA MOHAMAD. Advance Machine Learning Methods for Dyslexia Biomarker Detection: A Review of Implementation Details and Challenges

[3] Norah Dhafer Alqahtani, Bander Alzahrani, Muhammad Sher Ramzan. Deep Learning Applications for Dyslexia Prediction

[4] Yazeed Alkhurayyif, Abdul Rahaman Wahab Sait. Deep learning-driven dyslexia detection model using multi-modality data