A Convolutional Recurrent Neural Network-Based Model For Handwritten Text Recognition To Predict Dysgraphia

Neha N. Doshi
Department of Computer Engineering, DJSCE
Mumbai, India
nehadoshi291@gmail.com

Manali U. Maniyar
Department of Computer Engineering, DJSCE
Mumbai, India
manalimaniyar@gmail.com

Khushi K. Shah
Department of Computer Engineering, DJSCE
Mumbai, India
khushishah2604@gmail.com

Namrata D. Sarda
Department of Computer Engineering, DJSCE
Mumbai, India
namudsarda1234@gmail.com

Meera Narvekar
Department of Computer Engineering, DJSCE
Mumbai, India
meera.narvekar@djsce.ac.in

Debajyoti Mukhopadhyay
Department of Computer Engineering, WIDiCoReL Lab
Mumbai, India
debajyoti.mukhopadhyay@gmail.com

***Abstract* – Common learning disabilities include dyslexia, dysgraphia, dyscalculia, auditory processing disorder, and language processing disorder. About 10% of the school-going population suffers from one of the learning disabilities mentioned above. Screening of learning disabilities at the proper age is beneficial for the student and the parents. Despite constant advancement in technology, limited attention has been given to the screening and detection of these disabilities. The formal method of manual testing in hospitals is usually intimidating which influences the performance of the child thereby affecting the results. This work aims to automate the process of manual testing for dysgraphia. The model focuses on handwriting-based risk prediction of dysgraphia based on thoroughly researched machine learning techniques and algorithms. The system proposes a convolutional recurrent neural network-based model for handwritten text recognition that will contribute to dysgraphia screening. This model is not only helpful for parents but also for counselors, and teachers who come across multiple handwriting every day. It can help them analyze if they are doubtful of any child suffering from dysgraphia.**

***Keywords* –**
Dysgraphia, Dyslexia, handwriting, Optical Character Recognition (OCR), risk, prediction, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN)

I. INTRODUCTION

Learning disabilities are due to neurobiological factors. The functioning of the brain is affected in a way that affects cognitive processes related to basic skills such as reading, writing, or math. It is crucial to realize that learning disabilities can affect a person's life beyond academics and impact relationships with family, friends,and the workplace. Since difficulties with reading,writing, and math are recognizable problems during the school years, the signs and symptoms of learning disabilities are most likely to be diagnosed during this time. The range of normal variation in learning among children are large. Often parents fail to accurately differentiate whether their child's learning falls within the normal range or is indicative of a problem that requires intervention. Some individuals do not receive an evaluation until they reach post-secondary education or adulthood. When a learning disability is not detected early, diagnosed correctly, and treated effectively, it can cause several other problems. These additional difficulties may be emotional, and a child can show signs of sadness, frustration, or disappointment.

The proposed system aims at automating testing for dysgraphia. Dysgraphia is a learning disorder in which the person has trouble with writing. Experts view dysgraphia as a challenge with a set of skills known as transcription. These skills include handwriting, spelling, and typing. Dysgraphia results in unusual and distorted handwriting. Letter reversals, for example, b to d, letter substitutions, letter additions, and omissions are common mistakes made by people suffering from dysgraphia. Though technological advancements are made in all domains of medical sciences, less attention has been given to this domain. There is a lack of automated testing.

The purpose of this system is to use neural network-based handwriting recognition to detect lines written by children or people of any age group for that matter. A prediction algorithm is generated from validated conditions achieved after a thorough literature survey. Finally, this work aims to create a simple user interface that is more efficient and attractive to a child when compared to a paper-based test as currently prevalent for detecting dysgraphia formally. Such an automated risk prediction test will help raise awareness and increase technology-based testing for learning disabilities.

II. REVIEW OF LITERATURE

The system proposed by Ruchira Kariyawasam et al. [1] is a mobile game that aims to help in the diagnosis of common learning disabilities. It does both screening

2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS) | 979-8-3503-3583-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICISCOIS56541.2023.10100514

and intervention. Trained CNNs are used to detect spoken letters or words and written letters, words, or numbers. Outputs from the CNN are used as inputs to the models that help in the screening of learning disabilities. KNN algorithm is implemented for dyslexia prediction. To determine if a person has letter dysgraphia, an SVM classifier is used and for number dysgraphia, a random forest classifier is used. Dyscalculia detection quiz scores are also evaluated by making use of Support Vector Machine. High-accuracy models are obtained by making use of the respective machine learning algorithms. Although, there are certain limitations to the system - the data gathered for training the models is applicable to a limited age group. Apart from this, the application developed is only available in Sinhala thereby resulting in a language barrier.

The study by Zuzana Dankovičová et al. [2] addresses the problem of writing disability, specifically dysgraphia. This was accomplished using a variety of machine learning techniques, including random forest, support vector machine, and adaptive boosting. The dataset comprises 78 handwriting samples, and 52 handwriting attributes (e.g. velocity, acceleration, jerk, duration, pen lifts, etc.). Only children ages 10–13 (inclusive of 10 and 13) were included in the study. Principal component analysis was used in order to visualize attributes from handwriting, in two-dimensional space.

The paper by Katie Spoon et al. [3] proposes two new data sets of handwriting collected from children with and without dyslexia, and an automated early screening technique to be used in conjunction with current approaches, a multi-stream convolutional neural network to classify handwriting for suggestions and to accelerate the detection process. Datasets were collected using two approaches: Classroom collection, Parental collection. The Optical Character Recognition (OCR) software approach did not work well due to variability in handwriting. Modifications to existing handwriting detection approaches were made and these models were run on controlled datasets. Preprocessing was done by Arvanitopoulos & Susstrunk's seam carving technique and randomized patches were picked. Convolutional neural network (CNN) was applied to this task and implemented using Keras and TensorFlow. The network had 5 convolutional layers, 3 max-pooling (MP) layers, 2 fully connected (FC) layers and a dropout layer as an architecture. Though preliminary results are good, more data is required, and overfitting becomes the biggest challenge. In the future, authors try on implementing unsupervised approaches, like clustering, and plan to study in detail the features utilized by neural networks. This work might also be extended to new languages.

In a paper, by Iza Sazanita Isa et al. [4] a system is developed that analyses the handwriting of primary school children in order to detect the presence of dyslexia. The dataset comprises the handwriting of 30 dyslexic children who have written the lowercase letters b, c, f, p, and numbers 2, 5, 6, and 7. These characters are chosen as they are commonly confused by dyslexic children. The scanned dataset is preprocessed and then undergoes segmentation by the bounding box technique. Next, optical character recognition is used to extract the features of the characters present in the input image. The output obtained is the correct characters from the sample, i.e., the ones that match the initially chosen 8 characters. The automated accuracy is then calculated using this output and a manual accuracy is calculated as well. The overall accuracy obtained, that is the number of samples for which both accuracies were the same, is 73.33%. The level of dyslexia is classified as risk or low risk - risk for 0 - 4 correct characters and low risk for 5 - 8 correct characters.

The authors [5] suggest games designed to train visual spatial attention, rapid speech sound identification, and visual to speech mapping. The games are aimed to reduce the high therapy dropout rates. It uses PhoneGap to generate multiplatform software that works on phones, tablets, and the web. HTML cache manifest and local storage is used in the touchscreen-enabled application. The focus of this paper is limited to one game called paths which is a visual simulation game. It helps the child to improve his ability to analyze each element by rapidly engaging peripheral vision.

It is therefore observed that dysgraphia is a less researched topic as compared to other learning disabilities. A lot of work in this field uses restricted private datasets thus giving the model less flexibility to learn. Learning disabilities have been given high emphasis for children, although dysgraphia is observed in adults as well and awareness of the same must be increased. Most platforms focus on only one learning disorder. Since learning and attention problems usually coexist, a unified platform is required for such screening. In the proposed solution, we use a publicly available dataset that covers a wide variety of handwriting. Dataset used contains samples of the handwriting of people of all age groups not being restricted to just children. Thus, training the model for larger variations of dysgraphic handwriting.

III. SYSTEM METHODOLOGY

A. System Workflow

The student will be given a corpus from the customized dataset that contains sentences including the words that people with dysgraphia might face difficulties with, like b, d, w, m, p, q, u, n. The user will be given an audio clip randomly chosen from the above-mentioned dataset and is expected to write down what he/she hears in the audio. This handwritten text will be converted to document text using the handwritten text recognition model. The input for this model will be an image file of the child's handwriting. This image file will be pre-processed, and then it will be fed into the trained neural network model. The document text extracted will undergo the miscue analysis by comparing the obtained text with the corpus text to predict the risk of dysgraphia in students. System checks for the percentage error rate for these characters. If this error rate is greater than 50% then there is a probable risk of dysgraphia, and then the analysis and result is displayed [4].

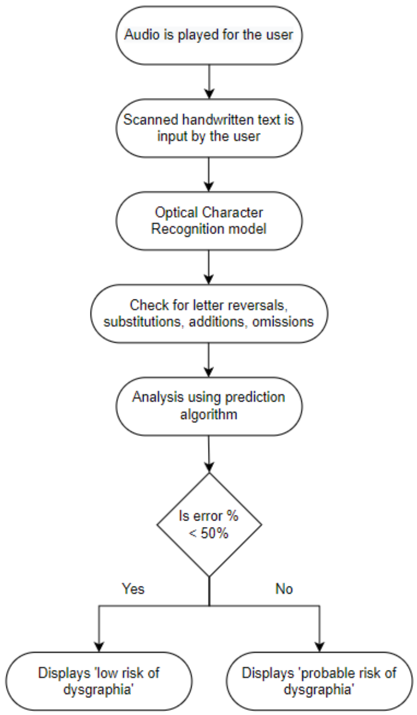

Fig 1. System Workflow

B. Dataset

The dataset used in this paper is the IAM dataset. Handwritten English text can be found in the IAM Handwriting Database, which is used to train as well as test the handwriting recognition model. Unconstrained handwritten text is recorded as PNG images with 256 gray levels in the database. There are 13353 photos of handwritten lines of text in the dataset, as well as 115320 isolated and labeled words recovered from scanned text pages using artificial segmentation. 657 writers provided examples of their handwriting to the dataset. Handwriting in the dataset is poorly spaced, slanting, and of various sizes. This ensures that the model will be able to receive higher accuracies when OCR is carried out on the answers given by test-attempters.

Fig 2. Handwriting dataset

C. Data Processing

The input images are first resized into 128 * 32px and then grey scaled. Then the image contrast is increased followed by the thickening of text lines by applying a morphological operation, thus making the text recognition more accurate. For photometric data augmentation, methods like gaussian blur, dilation, and erosion have been used, for better training and accuracy. For faster loading, we have used LMDB (Lightning Memory-Mapped Database) files.

D. Handwriting Recognition Model

The model for handwritten text recognition uses consists of 5 layers of convolutional neural network (CNN) 2 layers of recurring neural network (RNN) and a Connectionist Temporal Classification (CTC) loss and decoding layer. All 5 convolutional layers undergo 3 operations which are convolution operation, Non-linear ReLu function, and pooling operation. The other two layers are LSTM implementation of RNN. Here we have 80 classes representing 79 characters of IAM dataset and 1 CTC blank label. The CTC layer computes the loss value and decodes the input matrix into the final text.

E. Analysis Algorithm

The analysis algorithm receives its input from the handwriting detection model. It then checks for letter reversals, letter substitution, letter additions and omissions.
(a) For letter substitutions: if next of original = next of obtained, then count +1, else if, next +1 of original = next +1 of obtained, then count +2.
(b) For letter additions: if current of original = next of obtained, then count +1, else if, current of original = next +1 of obtained, then count +2.
(c) For letter omissions: if next of original = current of obtained, then count +1, else if, next +1 of original = current of obtained, then count +2.
(d) For letter reversals in word: if 'b','m','n','p, in current of original = 'd','w','u','q' in current of obtained or vice-versa, then count+1.
Along with this, the model also track the amount of time needed to complete the task. According to the general data the writing speed for following ages are:

- Age 6 = 3.6 wpm
- Age 7 = 5.6 wpm
- Age 8 = 7.2 wpm
- Age 9 = 9 wpm
- Age 10 = 10.4 wpm
- Age 11 = 12 wpm

Time taken more than the average rate would show the risk of having dysgraphia.

F. Selection of Threshold Value

A thorough literature survey of several papers that explore the medical domain of these disabilities was carried out. Each paper had slightly varying threshold values depending on the number of samples and method of testing. The system designed uses the threshold value of 50%. This has been derived from a paper by Iza Sazanita Isa et al. [4]. After additional validation from counsellors on the threshold value and considering that this threshold value is derived using machine learning techniques, 50% character error rate was found to be most suitable for this system.

G. Analysis and Risk Prediction

In the last step of the system workflow, the UI displays the thorough analysis and the risk factor. The system displays detailed and multiple analysis factors to make the output insightful for guardians, parents, and counsellors. Similarity between the corpus and obtained answer, number of mistakes and risk factor based on the character error rate is returned on the screen as the result.

IV. RESULTS AND DISCUSSIONS

The system uses a line-based handwriting recognition model and an analysis algorithm. On training the handwriting recognition model, an accuracy of 73/% was obtained for detecting words correctly. During validation, an additional measure of character error count was used for each word. Fig 3. shows the model validation and character error count for wrongly interpreted words from the validation dataset.

```
[OK] "the" -> "the"
[OK] "train" -> "train"
[OK] "." -> "."
[OK] "Catherine" -> "Catherine"
[OK] "is" -> "is"
[OK] "13" -> "13"
[OK] "." -> "."
[OK] "I" -> "I"
[OK] "suggested" -> "suggested"
[OK] "she" -> "she"
[OK] "might" -> "might"
[OK] "find" -> "find"
[OK] "them" -> "them"
[OK] "dif-" -> "dif-"
[OK] "ficult" -> "ficult"
[OK] "," -> ","
[OK] "but" -> "but"
[OK] "she" -> "she"
[OK] "said" -> "said"
[OK] "," -> ","
[ERR:1] "'" -> """
[OK] "Philip" -> "Philip"
[OK] "reads" -> "reads"
[OK] "them" -> "them"
[OK] "," -> ","
[OK] "doesn't" -> "doesn't"
[OK] "he" -> "he"
```

Fig3. Validation of the Model

Another accuracy measure used is character error rate with an early stopping of 10 during the training of the model. This ensures that the model does not overfit the training dataset. Fig 4. shows a graph of decreasing character rate and increasing accuracy aimed and obtained during the training of the model.

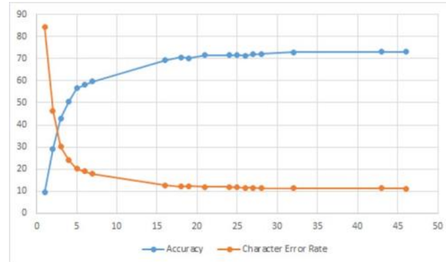

Fig 4. Accuracy and Character error rate over epochs

Multiple combinations of number of layers in the CNN and RNN were used to obtain a model with desirable accuracy. The internal workflow of the system plays an audio clip for the test attempter to write on paper. Sample testing was carried out on this model. Fig 5. shows the audio played and the sample image uploaded to the system. The output of the handwriting recognition model matches the text written by the test attempter. Higher accuracy without overfitting the dataset is aimed because the handwriting on children can vary over a wide spectrum of spacing, sizing and incline.

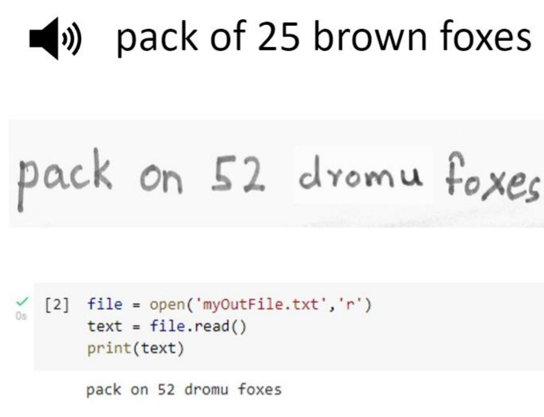

Fig 5. Sample testing of the OCR model

The analysis algorithm first calculates the number of correctly written words by the child. Then a similarity percentage base on LCS i.e. longest common subsequence is calculated. The number of letter substitutions, additions, omissions and reversals are calculated and displayed by the system next. The reason behind providing a detailed analysis along with the risk is to make the result as insightful as possible for a guardian or parent or a domain specialized counsellor guiding a child t attempt the test. This allows do decide the plan of action to help the child ease the difficulties he might be facing while writing words. Fif 6. Below shows the output on running the analysis algorithm on the sample test image provided in Fig 5.

The algorithm finally displays the risk of dysgraphia depending on the number of errors made in the 7 characters commonly mistaken and a threshold for errors to be 50%.

```
['pack', 'of', '25', 'brown', 'foxes']
['pack', 'on', '52', 'dromu', 'foxes']
similarity =  0.7727272727272727
[1, 0, 0, 0, 1]
correct word count =  2

Original Text: pack of 25 brown foxes
Detected Text: pack on 52 dromu foxes
-----------------------------------------------------------------
f n
2 5
5 2
b d
w m
n u

Letter Substitutions: 6

-----------------------------------------------------------------
b d 1
w m 1
n u 1

Letter Reversals: 3

-----------------------------------------------------------------

Total Errors: 6

-----------------------------------------------------------------
Characters that are difficult for children with dysgraphia: b, c, f, p, 2, 5, 6, 7
Total number of these characters in corpus text: 7
Total number of correct characters: 3
Note: Threshold is 50%
percentage correct characters =  42.857142857142854
probable risk of dysgraphia
```

Fig 6. Output of the analysis algorithm

The system aims to keep the user interface clean and simple. The test attempter can easily navigate through the interface to listen to the audio clip one is supposed to imitate on paper. The sample user interface of the result screen consists of the analysis and risk predicted on the sample testing image demonstrated in Fig 6.

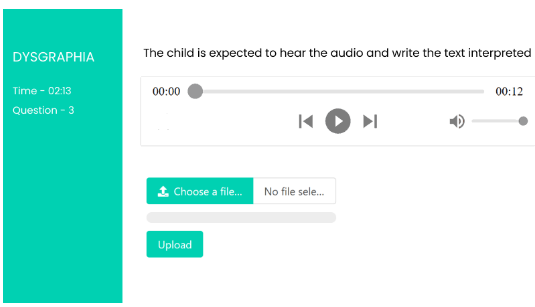

Fig 7. Test interface

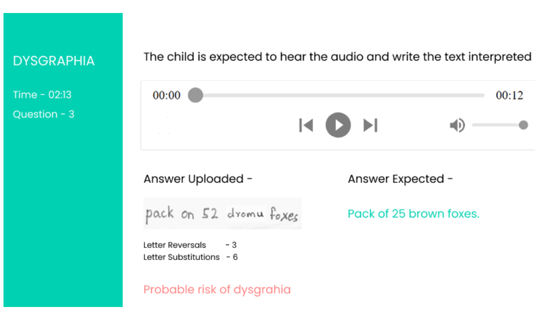

Fig 8. Sample Result UI

Fig 7 and Fig 8 demonstrate the proposed UI of the system. Thus, the results obtained by the system ensure easy automated testing along with insightful risk prediction of the learning disability of dysgraphia. The time parameter has shown promising results. The model continues to strive to ensure the results are accurate.

V. CONCLUSION & FUTURE WORK

The model proposed in this paper helps in detecting the risk factor for students with dysgraphia. It will calculate the risk factor on basis of the count of mistakes made by a student in imitating the corpus text. These mistakes include letter substitutions, letter addition, letter omission, and letter reversal in words. Since this model can also be used by parents, teachers, and counselors, it returns the detailed result of the 'total number of commonly misinterpreted letters by dysgraphic student' in corpus text, out of them the characters identified by the user correctly and the character error rate, for better understanding of the severity of user's condition. The threshold value for error percentage, to decide if the user has the risk of dysgraphia, is chosen after an extensive literature review in both the medical and technical domains. Using the described methods, the model was trained with 73% accuracy on the IAM dataset that includes the handwriting of dysgraphic children as well as adults to touch the corners of the dysgraphia condition to train the model.

Future work on the system can involve extending its capability to assess multiple learning disorders. Dyslexia, a similar learning disability related to reading can be easily incorporated into the system by switching the handwriting recognition model to a speech recognition model. Formal ways of testing can be replaced by accurately trained and tested systems like these making the entire process easier for children, parents, teachers, and counselors. The platform could be extended to have multiple native languages for a broader audience.

VI. REFERENCES

[1] R. Kariyawasam, M. Nadeeshani, T. Hamid, I. Subasinghe and P. Ratnayake, "A Gamified Approach for Screening and Intervention of Dyslexia, Dysgraphia and Dyscalculia," 2019 International Conference on Advancements in Computing (ICAC), 2019, pp. 156-161, doi: 10.1109/ICAC49085.2019.9103336.

[2] N. Giri et al., "Detection of Dyscalculia Using Machine Learning," 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 1-6, doi: 10.1109/ICCES48766.2020.9137871.

[3] Z. Dankovičová, J. Hurtuk and P. Feciľak, "Evaluation of Digitalized Handwriting for Dysgraphia Detection Using Random Forest Classification Method," 2019 IEEE 17th International Symposium on Intelligent Systems and Informatics (SISY), 2019, pp. 000149-000154, doi: 10.1109/SISY47553.2019.9111567.

[4] Katie Spoon, David Crandall, Katie Siek, "Towards Detecting Dyslexia in Children's Handwriting Using Neural Networks," ICML Workshop on AI for Social Good, 2019.

[5] Isa, Iza & Rahimi, Wan & Ramlan, Siti Azura & Sulaiman, Siti. (2019). Automated Detection of Dyslexia Symptom Based on Handwriting Image for Primary School Children. Procedia Computer Science. 163. 440-449. 10.1016/j.procs.2019.12.127.

[6] Facoetti, Andrea & Franceschini, Sandro & Gaggi, Ombretta & Galiazzo, G. & Gori, S. & Palazzi, Claudio & Ruffino, M.. (2014). Multiplatform games for Dyslexia identification in preschoolers. 1152-

1153. 10.1109/CCNC.2014.6940496.

[7] Asvestopoulou, Thomais & Manousaki, Victoria & Psistakis, Antonis & Andreadakis, Vassilios & Aslanides, Ioannis & Papadopouli, Maria. (2019). DysLexML: Screening Tool for Dyslexia Using Machine Learning. arXiv:1903.06274v1 [cs.CY] 14 Mar 2019.

[8] Richard, Gilles & Serrurier, Mathieu. Dyslexia and Dysgraphia prediction: A new machine learning approach. arXiv:2005.06401 [cs.OH], 15 April 2020.

[9] David, Julie & Balakrishnan, Kannan. Prediction of Learning Disabilities in School Age Children using SVM and Decision Tree, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (2) , 2011, 829-835.

[10] Gooch, Daniel & Vasalou, Asimina & Benton, Laura. (2015). Exploring the use of a Gamification Platform to Support Students with Dyslexia. 10.1109/IISA.2015.7388