Comparative Study of CNN Models on the Classification of Dyslexic Handwriting.

Subha Sreekumar
Dept.Computer Science and Engineering
National Institute of Technology Calicut
subha_p200120cs@nitc.ac.in

Lijiya A
Dept.Computer Science and Engineering
National Institute of Technology Calicut
lijiya@nitc.ac.in

***Abstract*—Developmental Dyslexia, one of the learning disabilities is a topic of scientific interest in a variety of disciplines such as psychology, speech and language therapy, data science, etc. While the reason for Dyslexia and its symptoms are still being researched by psychologists, data science is providing ways to intervene and detect them with the aid of technological advancements. Dyslexia is a neurological condition that impairs reading comprehension and has long-lasting impacts. But timely detection and intervention programs can alleviate its effects to a certain extent. This study aims to classify images of handwritten English characters into three classes namely: normal, corrected, and reversed, where normal class refers to normal handwriting, and corrected or reversed constitutes handwriting of children with Dyslexia. The dataset used for the study is available publicly on Kaggle. The building of an efficient CNN (Convolutional Neural Network) model for classifying dyslexic handwriting is the major emphasis of this work. This is accomplished by comparing several CNN models and evaluating how well they detect Dyslexia on the same dataset. The proposed CNN approach has demonstrated a sizable improvement in reliably classifying dyslexic handwritten images.**

***Index Terms*—Specific Learning Disability (SLD), Dyslexia, CNN , LeNet-5, Leaky ReLU.**

I. Introduction

Dyslexia is one of the Specific Learning Disabilities (SLD) along with others like Dysgraphia, Dyscalculia, and one or more combinations of them. The current definition of Dyslexia in the literature is "Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected about other cognitive abilities and the provision of effective classroom instruction" [1].

Regardless of above-average Intelligent Quotient (IQ) and healthy sensory systems, dyslexic children display problems related to spelling, fluency, accurate word recognition, as well as poor decoding abilities, all of which add up to a deficiency in the phonological aspect of a language [2]. Reading difficulty is the direct consequence of deficient phonetics and is considered the main symptom identifying Dyslexia. Since Dyslexia also manifests in a variety of other ways, researchers also examine other traits, such as handwriting, eye movement tracking, and brain activity monitored by an Electroencephalogram (EEG), among others.

Although difficult, early Dyslexia detection is essential for early intervention. Hence the need to identify symptoms as early as possible is urgent. Along with difficulty in reading, dyslexic children also have problems with automatic letter writing and naming which directly affects their ability to spell words [3]. In this aspect, this study focuses on the handwriting feature to identify possible symptoms of Dyslexia in children and compares the classification result with already existing work [4].

The work uses a publicly available dataset of handwriting images from Kaggle [5]. Classification of the dataset is carried out using the CNN (Convolutional Neural Network) model where CNN is used as both a feature extractor and a classifier. Deep Learning utilizes CNN for the recognition of objects in an image. CNN is a type of Artificial Neural Network (ANN) used for image recognition and classification. Among the different types of CNN models, LeNet-5 is simple and also the first of its kind. More importantly, LeNet was originally developed to categorize handwritten digits from 0–9 of the MNIST Dataset. Hence, in this study, the results of the proposed CNN model are compared with that of the LeNet -5 model to evaluate its performance.

II. Literature Review

Investigations into handwriting, speech, eye movement, brain activity, and games specially made for assessing different dyslexia-specific parameters are just a few of the studies related to Dyslexia identification. All these studies deal with different features that put together can shed light on the prediction of Dyslexia. Suggestions on the use of children's handwriting samples to identify possible indications of Dyslexia and explore modern machine learning methods for the purpose were made in [6].The work classified handwriting using neural networks to indicate the possibility of Dyslexia. CNN models have shown significant results in classifying the potential Dyslexia symptom based on handwritten images [4]. CNN with one convolution layer outperformed other models with a training accuracy of 0.985 but a validation accuracy of 0.86 which is about 10% less compared to the training accuracy. The LeNet-5 was modified in [18] to attain a test accuracy of 0.9534 and a test loss of 0.1583. Knowing that the

978-1-6654-9291-1/22/$31.00 ©2022 IEEE

network learns more effectively as it becomes deeper, or as the number of convolutional layers increases, further investigation of other CNN models was inevitable. Although there are not many research studies on handwriting for Dyslexia, studies on Dysgraphia, another learning disability that impairs the ability to write, do pay handwriting more attention [7], [8].

The use of audio recordings of words that are age-specific was recommended in [9]. since difficulty in reading is the primary symptom used to diagnose Dyslexia. Features like reading time and reading reaction time (interval between the initial display of the word and the start of reading) were computed and various machine learning algorithms were compared for the result. The neural network fared well against the others with 81.72% of accuracy.

Monitoring a user's eye movement while they read can reveal several cognitive processes that are going on. Individuals with Dyslexia have reading habits that are different from those of people without Dyslexia in several ways. This can be inferred by monitoring their eye movement patterns [10], [11].

Online gamified tests are gaining popularity as they make the detection of Dyslexia easier as well as encourage children to complete the assignment for detection. A machine learning model is used to predict Dyslexia using the data from game [1], [13]. The brain activities of dyslexics are studied through brain scans or Electroencephalography (EEG) along with machine learning classifiers to identify specific brain configurations [14], [15].

This study aims to work on the handwriting images to know if they could be of any help in the detection of Dyslexia symptoms and classify them as dyslexic and non-dyslexic using CNN deep learning models. Although Lenet-5 is an established CNN model for handwriting recognition, this study proposes a new CNN model and assesses its performance over LeNet-5 and its modification [18].

III. Methodology

A. Dataset

The study uses a publicly available dataset from the Kaggle database which was used in [4], [18]. It is a dataset containing images of handwritten English characters belonging to three classes of data: Normal, Corrected, and Reversed where both Corrected and Reversed are considered dyslexic and Normal class refers to the normal writing of the characters as suggested by Susan Barton of Bright Solution for Dyslexia. The training and test set consists of 208,372 samples in total, each of size 28x28 pixels. Although LeNet-5 was developed for handwritten digits, other CNN models are considered to evaluate their performance against LeNet-5. Examples of the dataset images are provided in Fig. 1.

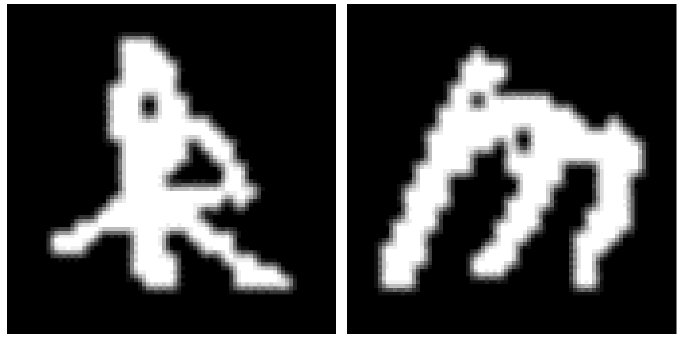

Fig. 1. Images of characters from dataset

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 24, 24, 6)         156
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 12, 12, 6)         0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 10, 10, 16)        880
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 8, 8, 16)          2320
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 4, 4, 16)          0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 2, 2, 120)         17400
_________________________________________________________________
flatten (Flatten)            (None, 480)               0
_________________________________________________________________
dense (Dense)                (None, 120)               57720
_________________________________________________________________
dropout (Dropout)            (None, 120)               0
_________________________________________________________________
dense_1 (Dense)              (None, 84)                10164
_________________________________________________________________
dense_2 (Dense)              (None, 3)                 255
=================================================================
Total params: 88,895
Trainable params: 88,895
Non-trainable params: 0
```

Fig. 2. Model summary of proposed CNN model

B. Proposed CNN Model

The model consists of 4 convolutional layers, 2 Max pooling layers, 3 Dense layers, and 1 Dropout layer. The first convolution layer takes a 28x28-sized image as input with a 3x3 filter. It captures low-level features such as gradient orientation, edges, color, etc. The other three convolution layers also have filters of 3x3 size each. The pooling layer after the convolution layers is Maximum pooling. The activation function used is ReLU (rectified linear unit) except for the last layer which has Softmax as an activation function. This model is trained using ReLU and also Leaky ReLU activation functions. Leaky ReLU is a variation of ReLU which fixes the "dying ReLU" problem by having a small slope for negative values instead of a flat slope and thus should perform better than ReLU. Also, the images were resized to 32x32 pixels to study the effect of convolutions on images of increased size.

The loss function used is categorical cross entropy and the optimizer used is Adams. Training is done for 20 epochs and a batch size of 60. The summary of the model obtained from the execution of the model on Kaggle is provided in figure Fig. 2. The visualization of the proposed CNN model is done using the Visualkeras Python package. It is given in Fig. 3. The proposed model is developed using Kaggle, an online community platform using Python language. Libraries such

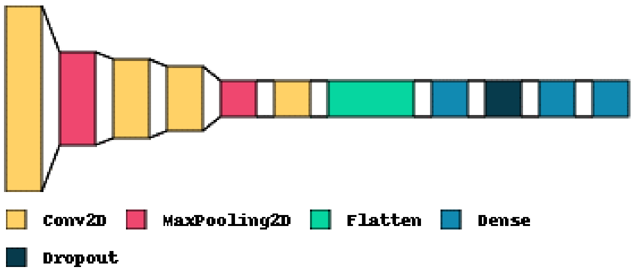

Fig. 3. Visualization of proposed CNN model

Fig. 4. Accuracy of proposed CNN model with ReLU

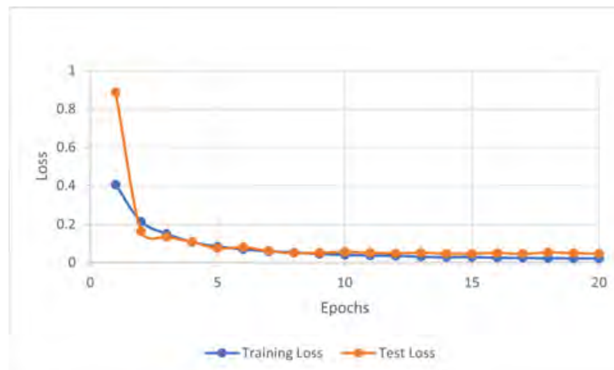

Fig. 5. Loss of proposed CNN model with ReLU

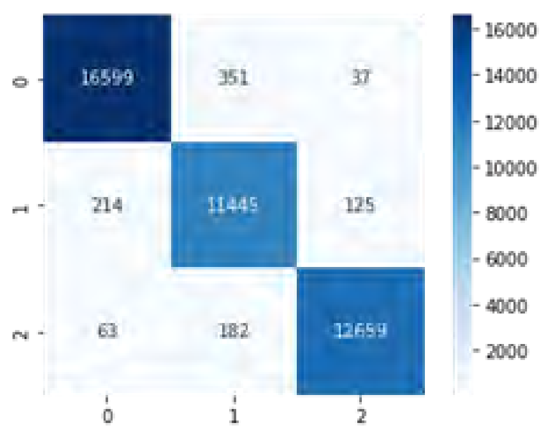

Fig. 6. Confusion matrix of proposed CNN model with ReLU

as TensorFlow and Keras are used to implement the Neural network.

IV. Results and Discussions

LeNet-5 has already been evaluated using the same dataset in [4], [18]. The findings of the proposed models are therefore compared to the results presented in these studies. Table 1 gives information about test accuracy and test loss for all the CNN models compared. The proposed CNN model with Leaky ReLU as an activation function has the best test accuracy among the studied models with 97.91%. But it is almost near 97.68%, the test accuracy of the proposed CNN model with ReLU as an activation function. Since no significant improvement is brought about by Leaky ReLU, the result of the proposed CNN model with the ReLU activation function alone is presented. The test accuracy results of models CNN-1, CNN-2, and CNN-3 as examined in [4], are 86%, 87%, and 86.5% respectively. The details of the models are explained in [4]. The test accuracies of LeNet-5 [18] and Modified LeNet-5 [18] are 88% and 95% respectively. These details are provided in TABLE I. From TABLE I, it is clear that the proposed model's performance is significantly better than other models that were examined. The Dropout layer in the model helps prevent overfitting of the model. Fig. 4 plots the accuracy of the proposed CNN with ReLU. Fig. 5 plots the training loss and test loss of the proposed CNN model with the ReLU activation function. The accuracy plot and loss plot both reveal that the model is not overfitting. TABLE II lists the classification metrics precision, recall, and F1-score of the classes. Since the values of the F1 scores for all three classes are closer to 1, it may be concluded that the model is more reliable. High values for them show that the proposed CNN model is efficient.

TABLE I
Performance Comparison

MODEL	TEST ACCURACY	TEST LOSS
CNN-1 [4]	0.86	0.55
CNN-2 [4]	0.87	0.5
CNN-3 [4]	0.865	0.6
LeNet-5 [16]	0.8873	0.2887
Modified LeNet-5 [16]	0.9534	0.1583
Proposed CNN-ReLU	0.9768	0.0827
Proposed CNN-Leaky ReLU	0.9791	0.0721

A. Confusion matrix analysis

Fig. 6 shows the confusion matrix of the proposed CNN model with ReLU. The axis shows 0 for the label Corrected,1 for the label Normal, and 2 for the label Reversed. It is evident that the model successfully identifies the classes of corrected and reversed writing that influences dyslexic handwriting.

B. 10-fold cross validation

To locate any potential overfitting, 10-fold cross-validation is carried out. It is stated that the mean accuracy is 0.98, the standard deviation is 0.194, and the mean loss is 0.071. Thus the proposed model is a generalized model.

TABLE II
PRECISION,RECALL AND F1-SCORE OF THE CLASSES

CLASS	PRECISION	RECALL	F1-SCORE
Corrected(0)	0.98	0.98	0.98
Normal(1)	0.96	0.97	0.97
Reversed(2)	0.99	0.98	0.98

V. CONCLUSION AND FUTURE WORK

This work reaffirms the use of handwriting to recognize a potential Dyslexia symptom using the CNN Deep Learning algorithm. Findings are favorable when comparing the proposed CNN model to the established LeNet-5 model. Thus this work successfully proposes a new CNN model that effectively classifies the 3 classes with an accuracy of 97.68%. It is also possible to design and test hybrid models [16], [17] that combine independent feature selection and classification models and determine whether they can improve the results. In the future, handwriting analysis of English words and even paragraphs can be used to automatically detect Dyslexia symptoms that are presently evaluated in person by professionals.

REFERENCES

[1] Lyon, G. Reid, Shaywitz, Sally E, Shaywitz, Bennett, “A Definition of Dyslexia.”,*Annals of Dyslexia*,v53 p1-14 2003.

[2] R. L. Peterson and B. F. Pennington,“Developmental Dyslexia”,*The Annual Review of Clinical Psychology*,vol. 11, no. 9, Jan., pp. 9.1-9.25, 2015.

[3] Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., Raskind, W. “Writing problems in developmental dyslexia: Under-recognized and under-treated.”,*Journal of School Psychology*,2008 Feb;46(1):1-21.

[4] I. Sazanita Isa, M. Aiman Zahir, S. Azura Ramlan, W.Li-Chih, and S. Noraini Sulaiman, “CNN comparisons models on Dyslexia handwriting classification.”,*ESTEEM Academic Journal*, Vol. 17, March 2021,12-25.

[5] I. Sazanita Isa, “Dyslexia Handwriting Dataset.”, Kaggle,[online], Available: https://www.kaggle.com/datasets/drizasazanitaisa/dyslexia-handwriting-dataset.

[6] K. Spoon, D. Crandall, and K. Siek, “Towards detecting Dyslexia in children’s handwriting using neural networks.”,*Proc. Int. Conf. Mach. Learn. AI Social Good Workshop*, 2019, pp. 1–5.

[7] J. Mekyska, M. Faundez-Zanuy, Z. Mzourek, Z. Galaz, Z. Smekal, and S. Rosenblum, “Identification and rating of developmental dysgraphia by handwriting analysis,” *IEEE Transactions on Human-Machine Systems*, vol. 47, pp. 235–248, 2017.

[8] S. Rosenblum and G. Dror, “Identifying developmental dysgraphia characteristics utilizing handwriting classification methods,” *IEEE Transactions on Human-Machine Systems*, vol. 47, pp. 293–298, 2017.

[9] J.Radford, G.Richard, H.Richard, M.Serrurier, “Detecting Dyslexia from audio record: An AI approach.”,*14th International Conference on Health Informatics*,Jan 2021.

[10] M. Modak, K. Ghotanae, S. V, N. Kelkar, A. Iyer, Prachi G., “Detection of Dyslexia using eye tracking measures.”,*International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, ISSN: 2278-3075, Volume-8, Issue- 6S4, April 2019.

[11] S. Shrestha, P. Murano, “An algorithm For automatically detecting Dyslexia on the fly”, *International Journal of Computer Science & Information Technology (IJCSIT)* , Vol 10, No 3, June 2018.

[12] L. Rello, R. Baeza-Yates, A. Ali, J. P. Bigham, and M. Serra, “Predicting risk of dyslexia with an online gamified test”,*PLoS One*, vol. 15, no. 12,2020.

[13] Rello, L., Romero, E., Rauschenberger, M., Ali, A., Williams, K., Bigham, J. P., and White, N. C., “Screening dyslexia for English using HCI measures and machine learning.”,*Proceedings of the 2018 International Conference on Digital Health, Lyon, France*, April 23-26, 2018, pages 80–84. ACM.

[14] Tamboer, P., H. C. M. Vorst, S. Ghebreab, and H. S. Scholte. ”Machine learning and dyslexia: Classification of individual structural neuro-imaging scans of students with and without dyslexia.”, *NeuroImage: Clinical 11* (2016): 508-514.

[15] Frid, A. and Manevitz, L.M,“ Features and Machine learning for correlating and classifying between brain areas and dyslexia.”,*arXiv preprint* ,arXiv:1812.10622,Dec 2018.

[16] Ali S, Li J, Pei Y, Aslam MS, Shaukat Z, Azeem M., “An effective and improved CNN-ELM classifier for handwritten digits recognition and classification. ”,*Symmetry*,2020; 12(10):1742.

[17] S. Ahlawat, A. Choudhary.,“Hybrid CNN-SVM classifier for handwritten digit recognition”, *Procedia Computer Science*, Volume 167, 2020, Pages 2554-2560.

[18] M. S. A. B. Rosli, I. S. Isa, S. A. Ramlan, S. N. Sulaiman and M. I. F. Maruzuki, “Development of CNN transfer learning for Dyslexia handwriting recognition,” *11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, 2021, pp. 194-199.