

# CNN-Based Dysgraphia Classification: A Comparative Study on Word and Paragraph-Level Handwriting Analysis

Elizabeth Vinitha P V <sup>1\*</sup>, Pramod Pavithran <sup>2</sup>

<sup>1,2</sup> School of Engineering, Cochin University of Science and Technology  
Kerala, India

**Aims:** To evaluate the effectiveness of ResNet50, MobileNetV2, and EfficientNetB7—for classifying dysgraphia-affected handwriting at the word and paragraph levels using a custom Malayalam dataset.

**Study design:** An experimental study using deep learning models for dysgraphia classification.

**Place and Duration of Study:** Conducted at Cochin University of Science and Technology, Kerala, over the research period of 2023-2025.

**Methodology:** The dataset consists of handwritten paragraph samples in the Malayalam language, collected from students with and without dysgraphia. A total of 143 samples were gathered, including 31 diagnosed cases of dysgraphia, with the guidance of professionals.

**Results:** MobileNetV2 achieved the highest accuracy of 90.91%, demonstrating superior performance due to its efficient feature extraction techniques.

**Conclusion:** Deep learning models, particularly MobileNetV2, show promising potential for dysgraphia classification, contributing to assistive handwriting analysis and early detection.

*Keywords: Dysgraphia classification, CNN handwriting analysis, Deep learning for learning disorders, MobileNetV2, ResNet50, EfficientNetB7, Dysgraphia detection.*

## 1. INTRODUCTION

Dysgraphia is a neurological disorder that primarily impairs handwriting and fine motor skills, presents significant challenges in educational and professional settings (Kedar et al., 2021). It can affect not only writing quality, but also the overall written product of writing like spelling, organization, letter size and word space. Several studies have been done and research have shown that approximately 10% to 30% of children worldwide experience difficulties with their handwriting ability. Diagnosing learning disorders like dysgraphia is particularly challenging as it requires analyzing multiple factors. Moreover, these symptoms of dysgraphia are varied and can differ based on a child's age and stage of development. Early identification of dysgraphia is important for implementing timely interventions and support strategies for children that can negatively affect their academic performance and overall quality of life (Drotár and Dobeš, 2020). However, the traditional assessment methods come with challenges. The test is completely relies on manual evaluation, which is both time-consuming and subjective. In manual dysgraphia assessments, a team of specialists primarily evaluates or reviews the child's handwritten product and assign a score based on their observations. Since this process can vary depending on the experts' experience and judgment. Then as a final step, diagnosis is typically confirmed by a clinical psychologist through further testing — a step that can add considerable time to the process (Kunhoth et al., 2024). To overcome these limitations, researchers have developed automated systems for diagnosing dysgraphia, highlighting the potential of deep learning (DL) algorithms to enhance handwritten text recognition (Dixit et al., 2020). Among these, Convolutional Neural Networks (CNNs) have shown remarkable success in image recognition tasks across a wide range of applications, such as medical image analysis and document processing (Tajbakhsh et al., 2016; Adamo et al., 2015).

On the basis of above, the present study aims to develop a novel CNN model to identify dysgraphia from handwritten images, distinguishing between two classes: dysgraphia-affected and normal. CNNs can automatically learn multi-level features from handwriting images and making them a promising tool for creating objective and efficient methods to detect dysgraphia through handwriting analysis (Harun et al., 2024). DL approaches, particularly CNNs, have shown significant potential in automating handwriting analysis. This study investigates the effectiveness of ResNet50, EfficientNetB7, and MobileNetV2 in classifying dysgraphia handwriting at word and paragraph levels.

This study aims to:

- Compare the effectiveness of ResNet50, MobileNetV2, and EfficientNetB7 in classifying dysgraphia-affected handwriting.
- Analyse the role of augmentation in improving model performance.
- Identify the best-performing model and its architectural advantages.

## 2. RELATED STUDY

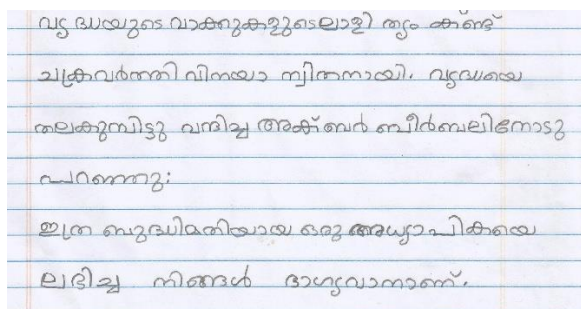
Detecting dysgraphia or other learning disabilities at an early age has been a key focus of researchers, with various methods being explored. Machine learning techniques have been widely used for classification tasks. A study by (Dutt & Ahuja, 2021) compared different machine learning classifiers, including K-Nearest Neighbors (KNN), XGBoost, Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machine (SVM), on a dysgraphia dataset, highlighting their strengths and limitations. Similarly, (Devi & Kavya, 2023) explored contemporary computational techniques for learning disability detection, providing an objective assessment of dyslexia. Another notable study by R. Kariyawasam et al. (Kariyawasam et al., 2019) developed an interactive game-based platform for assessment and remediation of dyslexia, letter dysgraphia, dyscalculia, and numeric dysgraphia., showcasing the integration of technology into early detection and intervention.

Handwriting recognition is one of the most resource-efficient and analytically straightforward methods for dysgraphia detection (Doshi et al., 2023). As a result, detecting dysgraphia through handwriting analysis has gained significant attention in recent years. DL methods, particularly CNNs, have emerged as the preferred approach for handwriting recognition, as they consistently yield significant results (Kunhoth et al., 2023; Kunhoth, Al-Maadeed, Saleh, & Akbari, 2024). However, CNN models often struggle with small datasets (Huang et al., 2020; Chakraborty et al., 2022). To mitigate this limitation, data augmentation techniques are widely used to generate synthetic images, enhancing CNN model effectiveness (Huang et al., 2020). Despite the general advantages of augmentation in improving deep learning models, its impact was minimal in our study. This suggests that the inherent features of dysgraphia-affected handwriting may require more advanced augmentation strategies or alternative feature enhancement techniques for better classification performance

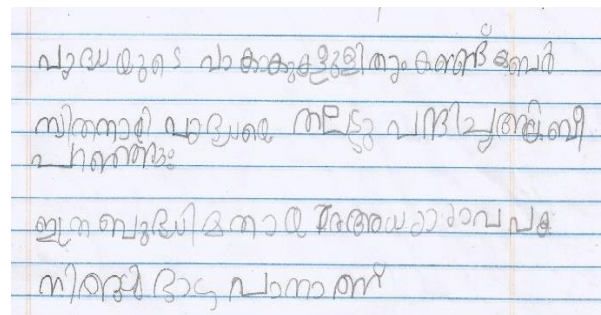
## 2. METHODOLOGY

### 2.1 Dataset

The dataset consists of handwritten words and paragraphs from dysgraphia and normal individuals. Dysgraphia samples were obtained from special schools with the assistance of two special educators, while normal samples were collected from mainstream schools with the help of class teachers. The dataset is categorized into two levels for analysis: word-level and paragraph-level handwriting.



(i)



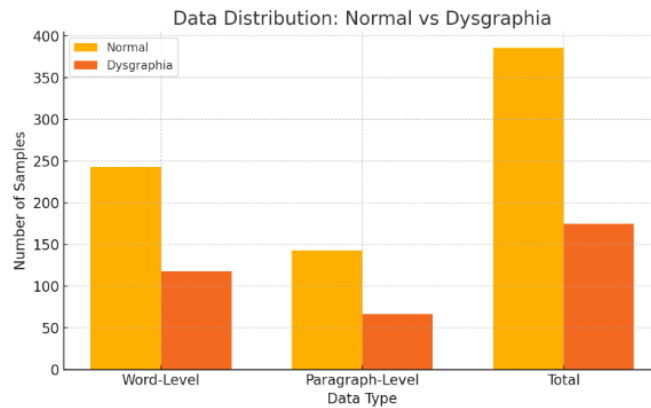
(ii)

**Fig. 1. Input Handwriting Samples (i) Normal (ii) Dysgraphia-affected**

Figures 1 presents examples of normal and dysgraphia handwriting samples from the dataset. Thus, these samples highlight the visual differences between the two categories names dysgraphia and normal writing. Each sample is labelled as either Dysgraphia or Normal for binary classification. Table 1 shows the data distribution, while Figure 2 provides its visual representation in the form of a bar diagram.

**TABLE 1. Data Distribution**

Data Type	Normal	Dysgraphia
Word-Level	243	118
Paragraph-Level	143	67
Total	386	175



**Fig. 2. Data Distribution**

## 2.2 Data Augmentation

Image augmentation is a widely used technique in DL that enhances a training dataset by applying transformations such as scaling, rotation, and flipping. This approach strengthens model robustness and generalization (Cimpoi, Maji, & Vedaldi, 2015) in classification. However, its effectiveness depends on the dataset characteristics and model architecture. In summary, data augmentation was applied to improve the model's adaptability to the training set. The following transformations were used:

- Rotation:  $\pm 30$  degrees
- Width & Height Shift:  $\pm 20\%$
- Shear Transformation: 20%
- Zoom:  $\pm 20\%$
- Horizontal Flip: Applied

By applying all data augmentation techniques together, the dataset size was effectively quadrupled. This enhanced diversity in the training data is anticipated to improve the performance and robustness of our deep learning models for dysgraphia identification in Malayalam datasets.

## 2.3 Deep Learning Models

Recently, tremendous advancements in DL, particularly in CNNs. It assured a significantly improved performance in texture classification tasks. These models, trained on large-scale image datasets like ImageNet and excel in feature extraction and pattern recognition. Thus, making them highly suitable for this task (Harika et al., 2024). Three DL architectures were used for classification:

### 2.1.1 MobileNetV2

The architecture of the proposed MobileNetV2-based model for classifying dysgraphia into normal and dysgraphia-affected categories is illustrated in Figure 3. The model starts with input images of handwriting, which are fed into the MobileNetV2 backbone pretrained on ImageNet. It passes through a series of convolutional layers and extracts relevant features. MobileNetV2 is a lightweight and efficient architecture that uses depthwise separable convolutions to reduce computational cost while maintaining the quality of feature extraction (Sandler et al., 2018). It uses several inverted residual blocks with linear bottlenecks, which progressively extract spatial and structural features from the handwriting. After feature extraction, a global average pooling layer reduces the spatial dimensions, which is then processed by dropout to prevent overfitting. Then a fully connected dense layer with 256 neurons and ReLU activation processes the features further, followed by another dropout layer. Finally, the output layer uses a softmax activation function outputs the classification results, the input into one of two categories: dysgraphia or normal. This architecture ensures efficient and accurate classification and is suitable for deployment in real-time or large-scale screening applications.

It is optimized for mobile and embedded vision tasks, provides an efficient trade-off between accuracy and computational efficiency. The architecture consists of:

- Base Model: MobileNetV2 (pre-trained on ImageNet)
- Global Average Pooling Layer
- Dense Layer: 256 neurons (ReLU activation)
- Dropout Layers: 50% and 30% to prevent overfitting
- Output Layer: Softmax activation for classification



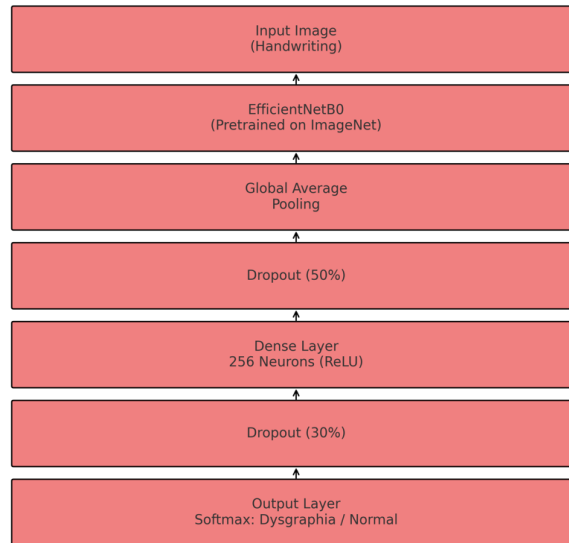
Fig. 3. Detailed Architecture of MobileNetV2-Based Model for Dysgraphia Classification

### 2.1.2 EfficientNetB0

The architecture of the proposed EfficientNetB0-based model for classifying dysgraphia into normal and dysgraphia-affected categories is illustrated in Figure 4. The model starts with input images of handwriting, which are fed into the EfficientNetB0 backbone pretrained on ImageNet. It passes through a series of convolutional layers and extracts relevant features. This model is a highly optimized CNN architecture that balances accuracy and efficiency by scaling its depth, width, and resolution. It is designed to capture hierarchical features while keeping computational cost low (He et al., 2016). Once the features are extracted by the base model, a global average pooling layer is applied. It reduces the spatial dimensions of the feature maps and helps to flatten the data without losing important information. Next, a dropout layer with a 50% rate helps prevent overfitting by randomly deactivating the neurons during training. A dense layer with 256 neurons and ReLU activation further processes the features, followed by an additional 30% dropout layer. The final softmax output layer performs classification into two categories: dysgraphia or normal. The architecture includes:

- Base Model: on EfficientNetB0 (pre-trained on ImageNet)
- Global Average Pooling Layer

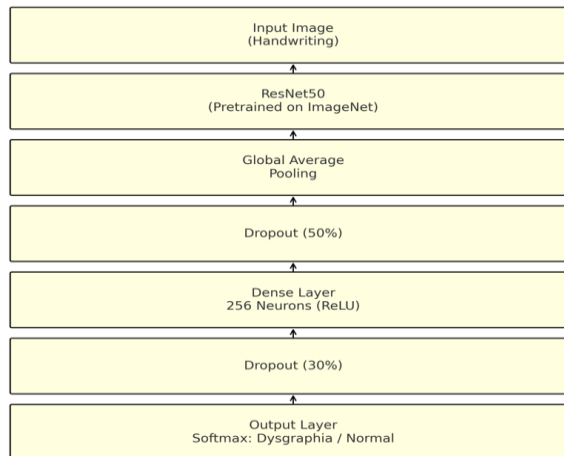
- Dense Layer: 256 neurons (ReLU activation)
- Dropout Layers: 50% and 30% to prevent overfitting
- Output Layer: Softmax activation for classification



**Fig.4. Detailed Architecture of EfficientNetB0-Based Model for Dysgraphia Classification**

### 2.1.3 ResNet50

The architecture of the proposed ResNet50-based model for classifying dysgraphia into normal and dysgraphia-affected categories is presented in Figure 5. The model accepts input images of handwriting, which are processed using ResNet50, a deep residual network pretrained on ImageNet. It passes through a series of convolutional layers and extracts relevant features. ResNet50 leverages residual learning through skip connections to enable the training of very deep networks without the problem of vanishing gradients, making it highly effective in extracting complex and abstract features. After feature extraction, a global average pooling layer compresses the spatial dimensions of the feature maps, followed by a 50% dropout layer to prevent overfitting. A dense layer with 256 neurons and ReLU activation then processes the feature representation, followed by an additional dropout layer (30%). Finally, a softmax activation function is used in the output layer to classify the handwriting as either dysgraphia-affected or normal. This architecture combines depth and robustness, making it suitable for accurate dysgraphia classification in real-world applications. It leverages a deep residual learning framework, enabling the training of significantly deeper networks compared to earlier models (Tan & Le, 2019). The model structure includes:



**Fig. 5. Detailed Architecture of ResNet50-Based Model for Dysgraphia Classification**

- Base Model: ResNet50 (pre-trained on ImageNet)
- Global Average Pooling Layer
- Dense Layer: 256 neurons (ReLU activation)
- Dropout Layers: 50% and 30%
- Output Layer: Softmax activation for classification

### 2.4 Training & Testing

The models were trained separately for word-level and paragraph-level handwriting classification. The dataset was divided into Training, Testing and Validation sets. The training process fine-tuned MobileNetV2, ResNet50, and EfficientNetB7 using transfer learning. It employed Categorical Cross-Entropy as the loss function and the Adam optimizer with a learning rate of 0.001. The models were trained with a batch size of 16 for 10 epochs, incorporating Dropout (0.3–0.5) to reduce overfitting. The models were tested using accuracy, precision and recall score to assess their performance.

### 2.5 Implementation details

The experiments were conducted using Google Colab with Python 3.8, TensorFlow 2.3.0, and Keras 2.4.3. The training and evaluation were performed on Colab’s cloud-based GPU environment to accelerate computation

## 3. RESULT

The performance of MobileNetV2, ResNet50, and EfficientNetB7 was evaluated on both word-level and paragraph-level handwriting classification tasks. The models were trained with and without data augmentation to analyse its impact on classification accuracy. The results are presented Table II, Table III and Table IV below.

**TABLE 2. Word-Level Classification Performance**

Model	Without Augmentation (%)	With Augmentation (%)
ResNet50	68.97	79.31
EfficientNetB7	68.97	68.97
MobileNetV2	72.41	82.76

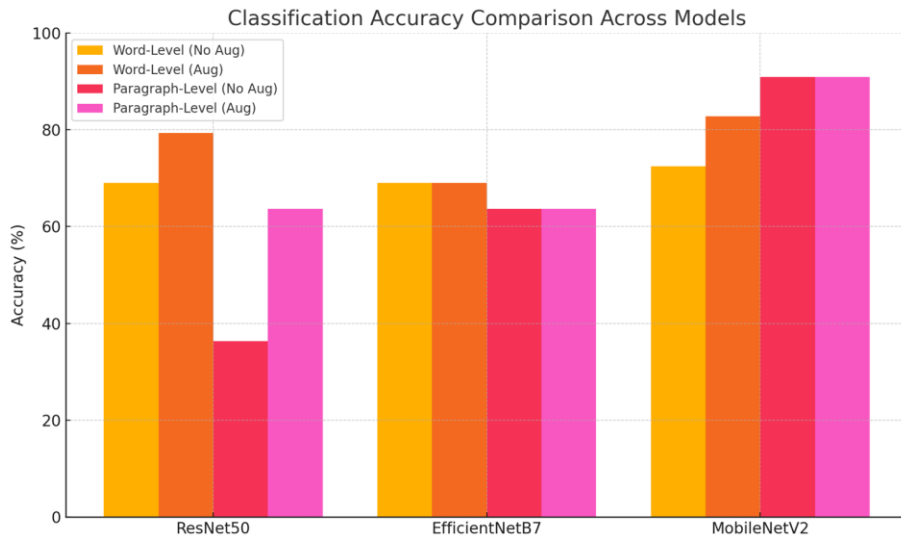
**TABLE 3. Paragraph-Level Classification Performance**

Model	Without Augmentation (%)	With Augmentation (%)
ResNet50	36.36	63.64
EfficientNetB7	63.64	63.64
MobileNetV2	90.91	90.91

**TABLE 4. Overall Performance Comparison**

Model	Word-Level Accuracy (%)	Paragraph-Level Accuracy (%)
ResNet50	79.31 (Augmented)	63.64 (Augmented)
EfficientNetB7	68.97 (Augmented)	63.64 (Augmented)
MobileNetV2	82.76	90.91

- 1) Word-Level Classification Performance:** As shown in Table II, MobileNetV2 achieved the highest accuracy of 82.76% when trained with data augmentation, followed by ResNet50 (79.31%) and EfficientNetB7 (68.97%). Without augmentation, the performance of all models dropped, highlighting the effectiveness of data augmentation in improving word-level handwriting classification.
- 2) Paragraph-Level Classification Performance:** For paragraph-level classification (Table III), MobileNetV2 significantly outperformed other models, achieving an accuracy of 90.91% in both augmented and non-augmented settings. ResNet50 demonstrated an improvement from 36.36% to 63.64% when trained with augmentation, while EfficientNetB7 remained consistent at 63.64%. The results indicate that MobileNetV2 is more robust in capturing paragraph-level handwriting features compared to the other architectures.
- 3) Overall Model Comparison:** A comparison of both classification tasks (Table IV) reveals that MobileNetV2 performed the best overall, particularly in paragraph-level classification (90.91% accuracy). ResNet50 showed notable improvement with augmentation, while EfficientNetB7 exhibited stable but lower performance across both levels. These findings suggest that lightweight models like MobileNetV2 are better suited for handwriting-based dysgraphia detection, benefiting from transfer learning and augmentation techniques.
- 4) Impact of Data Augmentation:** The results emphasize the positive impact of data augmentation, particularly for ResNet50 and MobileNetV2. In word-level classification, augmentation improved accuracy across all models, with an average increase of 6.9%. In paragraph-level classification, ResNet50 showed the most improvement (27.28% increase), whereas MobileNetV2 maintained a high accuracy regardless of augmentation.



**Fig. 6. Performance Comparison of Models with and without Augmentation**

The superior performance of MobileNetV2 suggests that lightweight architectures with depthwise separable convolutions are effective in extracting key handwriting features. ResNet50, despite its deeper structure, benefited significantly from augmentation, while EfficientNetB7, though highly parameterized, did not show notable improvements. These insights highlight the need for model selection based on computational efficiency and feature extraction capability, particularly when dealing with handwriting variability in dysgraphia diagnosis. Figure 2. shows the performance comparison with and without augmentation.

### 3. CONCLUSION

This study explored the classification of dysgraphia in Malayalam handwriting using MobileNetV2, ResNet50, and EfficientNetB7 at both word and paragraph levels. As far as we know, this is among the earliest studies to examine dysgraphia in Malayalam handwriting, addressing a significant gap in research. The results demonstrated that MobileNetV2 outperformed other models, achieving the highest accuracy, particularly in paragraph-level classification (90.91%). ResNet50 showed notable improvements with data augmentation, while EfficientNetB7 maintained stable but lower performance.

The findings highlight the effectiveness of lightweight architectures like MobileNetV2 for handwriting feature extraction in resource-constrained applications. Additionally, data augmentation played a crucial role in enhancing model performance, particularly for ResNet50. The study also underscores the challenges of paragraph-level handwriting classification, emphasizing the need for feature-rich models and improved preprocessing techniques.

Future research could explore alternative augmentation methods, transformer-based models, and explainable AI techniques to improve model interpretability and generalization. Expanding the dataset with more diverse Malayalam handwriting samples and integrating multi-modal approaches could further enhance dysgraphia diagnosis, paving the way for automated and language-inclusive handwriting assessment tools.

## **ACKNOWLEDGEMENTS**

Thanks to the special educators, teachers, and students who participated in the data collection process.

## **COMPETING INTERESTS**

Authors have declared that no competing interests exist.

## **REFERENCES**

- Kedar, S. V., et al. (2021). Identifying learning disability through digital handwriting analysis. *Turkish Journal of Computer and Mathematics Education*, 12(1S), 46–56.
- Drotár, P., & Dobeš, M. (2020). Dysgraphia detection through machine learning. *Scientific Reports*, 10(1), 21541.
- Kunhoth, J., et al. (2024). Automated systems for diagnosis of dysgraphia in children: A survey and novel framework. *International Journal on Document Analysis and Recognition (IJDAR)*, 27(4), 707–735.
- Dixit, U. D., et al. (2020). Recognition of handwritten word images. 2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC), 1–6. IEEE. <https://doi.org/10.1109/B-HTC50970.2020.9297818>
- Tajbakhsh, N., et al. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35(5), 1299–1312. <https://doi.org/10.1109/TMI.2016.2535302>
- Adamo, F., et al. (2015). An automatic document processing system for medical data extraction. *Measurement*, 61, 88–99.
- Harun, N., et al. (2024). Dysgraphia handwriting image augmentation for CNN model classification. 2024 IEEE 14th International Conference on Control System, Computing and Engineering (ICCSCE), 1–6. IEEE.
- Dutt, S., & Ahuja, N. J. (2021). Comparison of classification methods used in machine learning for dysgraphia identification. *Turkish Journal of Computer and Mathematics Education*, 12(11), 1886–1891.
- Devi, A., & Kavya, G. (2023). Dysgraphia disorder forecasting and classification technique using intelligent deep learning approaches. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 120, 110647. <https://doi.org/10.1016/j.pnpbbp.2023.110647>
- Kariyawasam, R., et al. (2019). A gamified approach for screening and intervention of dyslexia, dysgraphia, and dyscalculia. 2019 International Conference on Advancements in Computing (ICAC), 1–7. IEEE.
- Doshi, N. N., et al. (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), 1–6. IEEE.
- Kunhoth, J., et al. (2023). 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. (2024). Multimodal ensemble with conditional feature fusion for dysgraphia diagnosis in children from handwriting samples. arXiv preprint arXiv:2408.13754.
- Huang, Y., Jin, Y., Li, Y., & Lin, Z. (2020). Towards imbalanced image classification: A generative adversarial network ensemble learning method. *IEEE Access*, 8, 88399–88409. <https://doi.org/10.1109/ACCESS.2020.2993139>
- Chakraborty, P., Islam, A., Yousuf, M. A., Agarwal, R., & Choudhury, T. (2022). Bangla handwritten character recognition using convolutional neural network. *Lecture Notes in Data Engineering and Communication Technologies*, 132, 721–731.
- Huang, Y., Jin, Y., Li, Y., & Lin, Z. (2020). Towards imbalanced image classification: A generative adversarial network ensemble learning method. *IEEE Access*, 8, 88399–88409. <https://doi.org/10.1109/ACCESS.2020.2993139>
- Cimpoi, M., Maji, S., & Vedaldi, A. (2015). Deep filter banks for texture recognition and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3828–3836).
- Harika, K., Dhanalakshmi, P., Reddy, S. S. K., Shakeer, S., Thanmayi, M. S., & Haripriya, V. (2024, May). Handwritten signature recognition using MobileCNN. In *2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)* (pp. 1–7). IEEE.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4510–4520).
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).
- Tan, M., & Le, Q. (2019, May). EfficientNet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning* (pp. 6105–6114). PMLR.