

Detection of Dyslexia Through Images of Handwriting using Hybrid AI Approach

Norah Dhafer Alqahtani¹, Bander Alzahrani², Muhammad Sher Ramzan³

Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia^{1,2,3}
Information Systems, King Khaled University, Abha 61421, Saudi Arabia¹

Abstract—Dyslexia is a neurodevelopmental disorder characterized by difficulties with acquiring reading skills, despite the presence of appropriate learning opportunities, sufficient education, and a suitable sociocultural context. Dyslexia negatively affects children’s educational development and their acquisition of language, as well as their writing. Therefore, early detection of dyslexia is of great importance. The prediction of dyslexia through handwriting is an active research field of almost five years’ standing. In this paper, we propose hybrid models (CNN-SVM) and (CNN-RF) to reveal dyslexia through images of handwriting. The paper aimed to develop a CNN model to extract features from images of handwriting where CNN is highly reliable in extracting features from images, and to use SVM as a classifier due to its generalization abilities as well as using random forest (RF) as a classifier in (CNN-RF). The study aimed to combine a deep learning (DL) model and a machine learning (ML) model to improve model performance. Data sets that consisted of 176,673 images of handwriting were used in this study. The hyperparameter of the model was adjusted and examined in order to classify the three categories of handwriting. The CNN model that was built demonstrated an outstanding accuracy rate of 98.71% in effectively categorizing three distinct types of handwriting—99.33% with SVM, and 98.44% in the CNN-RF model. The aim of recognizing dyslexic handwriting through CNN-SVM was successfully attained, and our model outperformed all previous models.

Keywords—Dyslexia; dyslexia detection; deep learning; dyslexia classification; CNN model; SVM model; random forest model

I. INTRODUCTION

Dyslexia is a language-based, neurobiological, developmental learning disorder that affects how people learn to read (in terms of accuracy and speed) and how they learn to spell. As a result of impairments in the phonological component of language, individuals with dyslexia struggle to connect spoken language with the written word [1]. Difficulty accurately and fluently deciphering words can impact reading and vocabulary development [2].

Although dyslexia is neither a sign nor an indicator of low intelligence, it may cause a person to perform poorly academically and to become frustrated. A person may drop out of education entirely. Detecting dyslexia in children as early as possible and using assistance tools and intervention programs may improve their skills and learning performance. In recent years, a greater awareness of dyslexia and other learning difficulties has developed.

Traditional diagnostic tools and tests have been used to detect these disorders and assist individuals based on the results of diagnostic tests. These detection techniques have focused on behavioural elements, such as proficiency in reading and writing and working memory and have also included IQ tests.

They are generally standardized tests [3]. Since every individual with dyslexia has a unique experience of the disorder, this approach can be time-consuming and runs the risk of missing cases or providing inaccurate diagnoses. Traditional machine learning methods and deep learning algorithms are being increasingly used in dyslexia and biomarker detection and have proven their effectiveness in making diagnoses. One of the deep learning models is Convolutional Neural Network (CNN), which has the ability to deal with images and extract features from them. This paper aims to enrich research in this field; our study is considered the sixth study in predicting dyslexia by looking at handwriting using DL models. In addition, the study developed CNN model to raise its ability in extracting features. Moreover, It builds hybrid models based on CNN; (CNN-SVM and CNN-RF) for developing dyslexia detection through the use of machine learning (ML) and deep learning (DL). Better classification accuracy can be achieved when the DL and ML models are combined than when they are used separately, as noted in [4], where the first (DL) can automatically extract useful spatial characteristics, while the second (ML) can classify the features that have been extracted by DL mode (CNN). The paper is organized into five sections: Section II details the related work, noting studies that have examined the detection of dyslexia through children’s handwriting and their key results. Section III introduces the phases that will be followed in this study, Section IV presents suggested models that will be used in this study. Section V illustrates the steps of the experiment, starting from data set acquisition to classification phase. Section VI explores the results of this study. Section VII is the conclusion.

II. RELATED WORK

The advent of AI and its various capabilities have opened up possibilities for automating methods of diagnosis and early detection of dyslexia through the use of ML and DL. Existing methods of detecting dyslexia have included detection using Magnetic Resonance Imaging (MRI), Electroencephalogram (EEG) signals, and eye-tracking, as well as Electrooculogram (EOG) signals. The ML method and computer vision can help with dyslexia classification as illustrated in [5]. Although earlier methods achieved good accuracy, it was time-consuming and expensive to collect the required neurological data. Dyslexic children have irregular handwriting and reverse their letters. This feature (inverted letters) helped Spoon et al. [3] to look at a new method for dyslexia prediction—the use of images of handwriting. They attempted to collect samples to investigate if there was a possibility of diagnosing dyslexia through handwriting samples using the CNN model. They achieved 55.71.4% accuracy on average—better than

the random standard of 50%. These preliminary outcomes nevertheless showed promise, with the detection rate much higher than that able to be achieved by instructors and parents at that time. These results encouraged researchers and led them to prove this concept in their study [6]. They used 100 samples of handwriting and built CNN models for prediction. The result was 77.6%—higher than the previous result. In another study, Isa et al. developed an automated handwriting recognition framework using an image-processing and pattern-recognition technique through the use of MATLAB to build an Artificial Neural Network (ANN) model—a method for processing information that was inspired by how biological neural systems process information [7]. The study employed four letters—‘f,’ ‘p,’ ‘b,’ and ‘c’—as well as four numbers—‘5,’ ‘2,’ ‘6,’ and ‘7.’ These samples were chosen because dyslexic people are often confused about the shape of these figures. The samples were used to train and test a suggested model—ANN—to extract features from images and utilized MLP as a features classifier. The model achieved accuracy of 73.33%, which was still considered low, due to the lack of samples. This study was followed by an Indian study that aimed to determine whether it was possible to detect dyslexia in Indian handwriting [8]. The researchers were able to collect approximately 267 pictures from the participants’ textbooks and process them. The CNN model was used to extract features from the images and predict dyslexia in children; here it was able to automatically find strong features by using Keras and TensorFlow—the accuracy average was 86.14%. All previous studies suffered from the small-scale size of the data sets and the problem of an imbalanced class, which affected the performance of the DL models as well as well as their tend to the large-scale class. In the study [9] researchers put more effort into building a dyslexia data set that could be utilized by other researchers in the future, they used these data sets to build different CNN models to compare the performance of these models. They employed data augmentation techniques to maximize the scale of data sets and solve the class imbalance problem. Different CNN models, like CNN-1, CNN-2, CNN-3, and LeNet-5, were used and compared in this study. The CNN-1 model achieved high accuracy compared with other models, which reached 87% in the classification process. The transfer learning technique was used to develop the performance of the DL model for classification purposes in [1]. The study utilized the same data set as the previous study and built a CNN model based on the well-known handwriting detection architecture of LeNet-5. The suggested model achieved an outstanding accuracy of 95.34% in the three classes of classification. These studies have some limitations which illustrated in Table I.

This study presents the following contributions: (i) Develop the CNN model to raise its performance in feature extraction from the handwriting model; (ii) assessing the performance of SVM and RF classifiers, as well as a CNN classifier, in terms of their ability to utilize CNN feature extraction for improving diagnosis accuracy. These models will use the dataset that has been used by the last two studies after removing the redundancy from it, and the classes will be balanced.

III. METHODOLOGY

The study passes through three phases: the first phase relates to data preprocessing and dividing the dataset (training, validation, and testing) before used by the CNN model. In the

second stage, the CNN model will be built and used a dataset that has been processed in phase 1 (training, validation) to train the model in extracting features from the handwriting images till reach the required accuracy. In the last stage, training and validation data will be used as one data set to train the different classifiers, then test data will be used to test the performance of different classifiers. Fig. 1 The workflow of the proposed method. illustrates the workflow of this study. The primary objective of the research is to introduce and evaluate hybrid models, specifically (CNN-SVM) and (CNN-RF), for the effective detection of dyslexia through images of handwriting. The research leverages Convolutional Neural Networks (CNN) to extract meaningful features from the handwriting images, exploiting the well-known reliability of CNNs in image feature extraction. In addition, Support Vector Machines (SVM) are chosen as classifiers due to their strong generalization capabilities. The (CNN-RF) model introduces Random Forest (RF) as an alternative classifier, enriching the model diversity. This hybrid approach combines deep learning (DL) models, such as CNN, with machine learning (ML) models like SVM and RF, with the intention of enhancing overall model performance.

IV. SUGGESTED APPROACHES

The study suggested hybrid models based on CNN model for classifying data in this study (CNN-SVM and CNN-RF). In addition, developing CNN model that have been used in the literature to raise its ability in extracting features and the raise the performance in the prediction of dyslexia.

A. Convolutional Neural Network

The CNN model is a DL technique that is specifically designed for the task of image classification. It operates on two-dimensional (2D) images as input data. Similar to ANNs, the CNN model exhibits a hierarchical architecture consisting of multiple layers, where the output of each preceding layer is systematically linked to the input of the subsequent layer. A traditional ANN structure is paired with a stage for extracting spatial features using a sequence of convolutional filters [4]. The architectural design of the CNN model typically encompasses three primary, interconnected layers—namely, the convolutional layer, pooling layer, and fully connected layer. The initial step in the convolutional layer involves the computation of weights through the application of a convolution filter. This filter performs a dot product operation on either the 2D input data or the outputs of preceding layers within a localized region [4]. Feature maps are generated by using a nonlinear activation function, such as a Sigmoid, Tanh and Rectified Linear Unit (ReLU), which is what was utilized in this study. The pooling layer is utilized to condense the retrieved features into representative values, such as maximum or mean values, in order to simplify the information. The utilization of the max pooling layer has been extensively employed in the classification process of CNNs [10]. The two layers (convolutional and pooling) are sequentially arranged in an alternating manner until higher-level features are obtained. Once the convolutional and pooling processes have been performed to extract the high-level features, the resulting feature maps are converted into a one-dimensional vector and subsequently passed to the fully connected layer. Typically, the final, fully connected

TABLE I. LIMITATIONS OF THE RELATED WORKS

Study	Limitation
[3]	1) They achieved 55.7% accuracy, and the baseline was 50%, there was no significant improvement. 2) To improve the outcomes, more data are needed, particularly data from dyslexic pupils.
[6]	1) They achieved 77.6% accuracy results, still, there is a chance to improve it. 2) Their dataset consists of 100 samples, which are considered small-scale data. CNN does not perform well with small-scale data.
[7]	1) The data set is small and focused on specific small letters 'b', 'c', 'f', 'p'. 2) The performance of the classification accuracy does not exceed 75%. This is because the ANN needs a lot of samples to get high.
[8]	1) In this study, the letters are cropped manually, which required improvement, Handwritten recognition by using (OCR) can be used, or cursive and skew methods. 2) The dataset used in the experiments is small. More data is required to study the results further.
[9]	1) The imbalance class problem is still in this study although data augmentation techniques have been applied, in addition to the presence of some duplicate images in the dataset. 2) The study requires more dyslexic handwriting images for the test set (real dataset) in examining the performance of the model.
[10]	1) The imbalance class problem is still in this study although data augmentation techniques have been applied, in addition to the presence of some duplicate images in the dataset. 2) Other data augmentation techniques can be applied where the study focused on the rotation technique. The transfer learning effectiveness from sources was not explained well. 3) It can be improved classification tasks by using other methods and rise the model performance.

layer of a neural network is responsible for normalizing the network's output. This normalization process involves utilizing a SoftMax function to get probability values corresponding to the expected output classes. Ultimately, the categorization outcome is determined through the use of the maximum probability rule [4]. CNN has produced outstanding results in the field of computer vision and pattern recognition [11], for example in visual recognition [12], image retrieval [13], and scene annotation [14]. This model has been effectively implemented for character recognition in handwritten images. It has been successfully used in offline, handwritten Javanese character recognition [15]. In the study [12], CNN achieved 88% in the Arabic data set, and high accuracy in MODI (ancient Indian script) character recognition [16]. According to [17], modifying the CNN with two different types of training input—reconstruction feedback and classification feedback—was able to achieve an accuracy rate of 99.59% on the MNIST data set. Some of the research tweaks CNN in a number of ways to improve its performance and accuracy rate. Other research alters the input data to enhance the accuracy rate of the CNN model.

B. Support Vector Machine

SVM is a supervised learning model with corresponding learning algorithms that analyze data utilized for classification and analysis of regression. This model is considered one of the most solid prediction methods, which depend on statistical learning frameworks. It can be utilized to solve different problems in the real world; for example, it is useful in the categorization of text and hypertext, the recognition of handwritten characters [18], face detection, and satellite data classification. There are two kinds of SVM: linear and nonlinear. The first is utilized for linearly separable data. A data set is said to be linearly separable if it can be divided into two classes using a single straight line; the classifier used is known as a linear SVM classifier. However, if the data set cannot be classified using a straight line, such data is known as nonlinear data, and the classifier applied is called nonlinear. The basic idea of SVM is to get the best hyperplane, which maximizes the hyperplane margin. A good generalization is achieved by a hyperplane with a maximum margin.

CNN is an extension model of multilayer perceptrons (MLP), since its theoretical learning technique is the same as MLPs'. The MLP learning algorithm tries to reduce errors in the training set: it depends on empirical risk minimization (ERM). When the backpropagation method discovers the first separating hyperplane, whether it is a global or local minima, the training operation ends. The MLP learning algorithm does not continue to enhance the separating hyperplane solution. Further, the SVM classifier utilizes a structural risk minimization (SRM) precept on unseen data to minimize the errors of generalization, with a constant distribution for the training set [19]. Therefore, SVM generalization ability is much better than that of MLP. According to [20]–[23], the SVM method has high generalization performance, which means it can correctly classify data that has never been seen before. One paper [24] advised using the SVM classifier as an end classifier because it has better generalization ability than neural networks on standard CNN.

C. Random Forest

RF is widely utilised in the field of machine learning and falls under the category of supervised learning techniques. ML can utilise this technique for both regression and classification tasks. The approach is rooted in the principle of ensemble learning, wherein many classifiers are integrated to address intricate problems and enhance the model's performance [25]. The classifier known as "Random Forest" is so named because it consists of many decision trees that are constructed using different subsets of the provided dataset. By taking the average of the predictions made by these decision trees, the RF classifier aims to enhance the accuracy of its predictions for the supplied dataset. The random forest algorithm utilises an ensemble approach by aggregating predictions from several decision trees. By considering the majority vote among these predictions, the random forest algorithm generates the final output. We chose these classifiers for this study for two reasons: previous studies relied on this CNN classifier to classify dyslexia, and its use here after its development shows an increase in the performance of the model compared to the previous one. In addition, RF is considered an effective and robust algorithm due to the avoidance of over-fitting, its high level of accuracy of classification, its ability to assess the

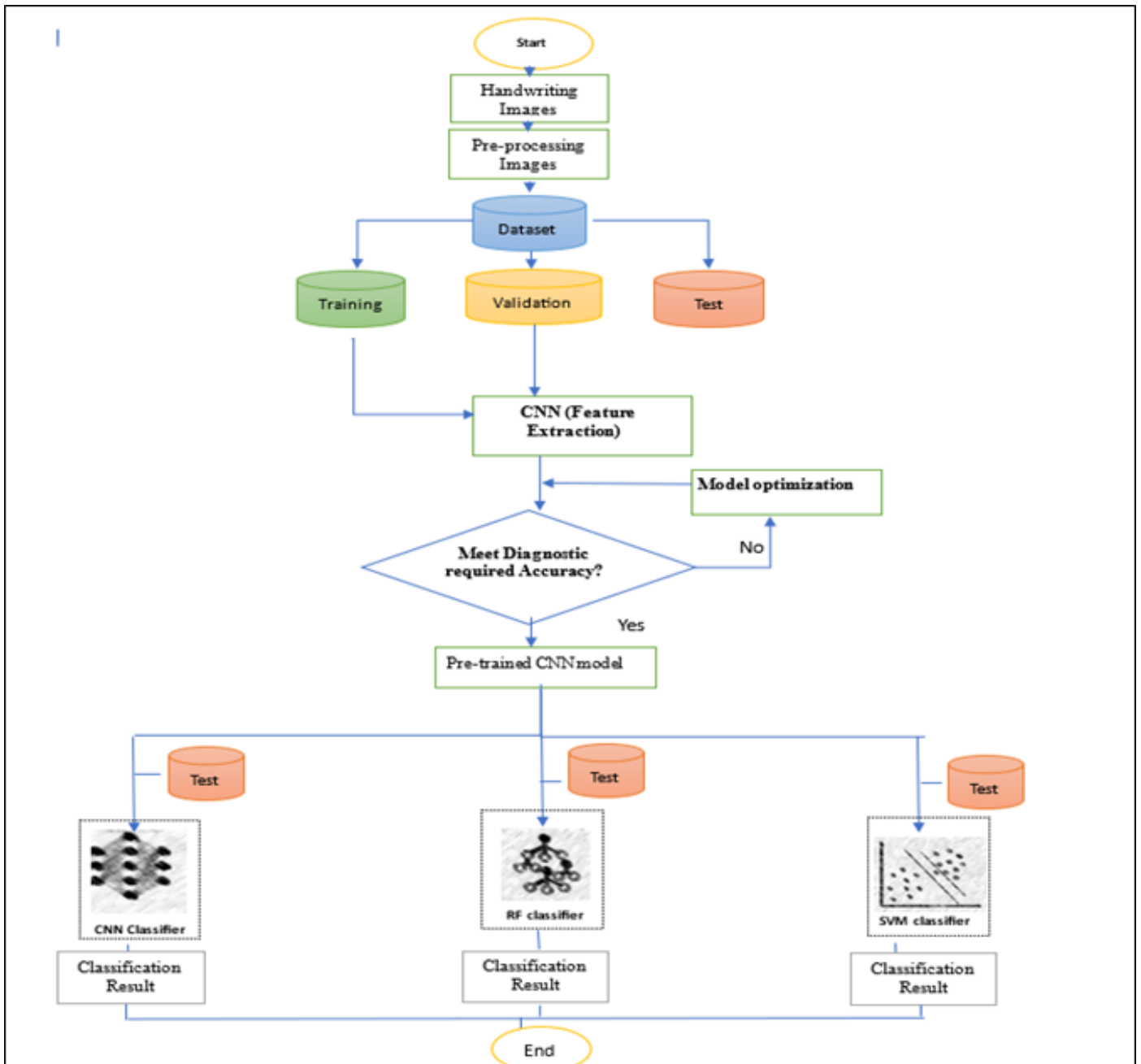


Fig. 1. The workflow of the proposed method.

relevance of variables [26] and its ability to operate effectively on huge databases [27] as well as the ability of SVM in generalization as we explained before.

V. EXPERIMENT

A. Data Set Used

The data set used in this study was images of the handwriting of three classes: reversal (dyslexics), normal, and corrected handwriting suggested by Susan Barton, founder of Bright Solutions for Dyslexia [28]. This data is publicly available in [29] which is obtained from three distinct sources. The uppercase letters were sourced from the NIST Special

Database 19 [30] while the lowercase letters were obtained from the Kaggle data set [31]. Additionally, certain data sets for testing purposes were collected from dyslexic students attending Seberang Jaya primary school by researchers in [9]. This data set consisted of 176,673 images. In the context of the reversal class, the normal handwriting data set underwent a mirroring process, resulting in the creation of reversal data sets through horizontal flipping as mentioned in study [9] as well as apply rotation and noise injection techniques after horizontal flipping. This data set has been used by the last two studies in this field [9], [1]. Samples of this data set are shown in Fig. 2 Samples of datasets used., where A is a sample of the reversal class, B is a sample of the normal class, and C is sample of

corrected writing.

B. Pre-processing

The data sets needed to be processed before feeding the classifier by them. A foreground-background swap was adopted to reduce computational overhead, as training an image with more white points (value 1) than black points (value 0) consumes more power and memory [32]. This procedure alters the background color to black while leaving the handwriting white. The subsequent action is the process of cropping the image to isolate the section containing the written content. This procedure involves the removal of undesired sections of an image, specifically from the bottom and top, and the right and left sides. The resulting image will be centered on the alphabet, thereby emphasizing it. In our study, the photos underwent a resizing process to dimensions of 32x32 pixels, ensuring uniformity across all data sets. This was done to facilitate their use as input for the CNN model. Ultimately, the entire data set was converted into a .csv file based on the one-hot encoding technique. The data sets were divided into 70% handwriting data sets for training objectives, 15% for validation, and 15% for testing objectives.

C. Feature Extraction

Feature extraction from the handwriting data sets was done through the CNN layers, which consisted of four convolutional layers, two max-pooling layers, and a flattened layer. Fig. 2 shows these layers. Convolutional layers act as a feature extractor; they receive the feature representations of the input pictures, and the trainable convolutional kernel adjusts its kernel weights automatically during the backpropagation training process [33], [34]. The pooling layer works to alter the input feature into a statistical picture of the surrounding feature, hence making the next feature smaller than the one before it [34]. As shown in Fig. 3, a batch normalization layer subsequent to each convolutional layer is suggested. This is used in neural networks to normalize the activation values of hidden units. Normalization ensures that the activations maintain consistent behavior during training, leading to improved accuracy and faster training [35]. Moreover, to prevent overfitting of the model, we utilized a dropout layer, which implements a regulatory mechanism wherein a subset of neurons is randomly disregarded during the training process. After being built, the model required fine tuning to raise performance and therefore accuracy in the training data sets. According to our objectives in the three class classifications to predict dyslexia, we preferred to utilize the ReLU as an activation function, as it addresses the issue of gradient vanishing that arises due to the utilization of sigmoid and tanh activation functions in deep neural networks [36]. Moreover, it enhances the intricacy of the neural network by incorporating non-linearity, hence enabling the network to acquire more intricate representations of the data, thereby increasing the performance of the model. The features extraction vector has generated from the convolution operation, the pooling operation as well ReLU function, as shown below equation,

$$O_{x,x} = \sum_x \sum_y f[l, k] * [x - l][y - k] \quad (1)$$
$$F_{x,y} = \max(0, F_{O_{x,y}})$$

The handwriting images are fed via CNN layers starting from convolutional layer for extracting significant features. The given input consists of a two-dimensional matrix with a rank of 2. The matrix has M rows and N columns, where the indices for the rows and columns are denoted as (x, y). It is important to note that the values of x and y should satisfy the condition $0 \leq x \leq M$ and $0 \leq y \leq N$. The convolutional operation layer produces the final feature map values, denoted as $F_{x,y}$ which are deemed to be significant for the task at hand. The utilization of an activation function is implemented at every layer in order to enable the model to effectively address nonlinear problems, as demonstrated in Eq. 1. Additionally, the incorporation of dropout and max pooling techniques serves to reduce the computing burden associated with the model.

The Adam optimizer was also used; the Adam optimization technique has gained wider acceptance in recent years for its application in deep learning tasks related to computer vision and natural language processing. It combines the advantages of AdaGrad and the RMSProp optimizer. One of the setup parameters utilized in the Adam optimization algorithm is the learning rate, which is assigned a value of 0.001. The last feature map undergoes a transformation so as to be represented as a vector with a single column. The recognition experiment involves feeding a single-column feature vector, which consists of identifiable features, to the classifier (a soft max layer of CNN, SVM, and RF) as shown in Fig. 3.

D. Classification

Following the completion of the pre-processing and feature extraction procedures, we classified the handwritten digit images using the different classifiers. Different classifiers were trained by utilizing feature vectors stored in a matrix format. The evaluation of the numerical value was conducted using the outcomes obtained from the training process. The hyperparameter of SVM employed an RBF function as its kernel, the cost parameter $C = 1$ and degree = 3. The RF was used to evaluate the model classification with hyperparameter (n_estimators), which was equal to 50 in this study, and to evaluate classifications with a fully connected layer in the CNN with a batch size = 64 and number of epochs = 40.

VI. RESULTS DISCUSSION AND OBSERVATIONS

In this section we evaluate the performance of the three models: the CNN model, the CNN-RF model, and the CNN-SVM model.

Loss and accuracy: The evaluation of a CNN model's performance in extracting features is determined by measuring the loss and accuracy on both the training and validation data sets. The testing dataset will be used in classification stage after the training all classifier as we explained in methodology section. The analysis and interpretation of this loss value provide insights into the model's effectiveness on these two sets. The error aggregation is computed as the cumulative aggregate of errors made for each individual example within the training or validation sets. The loss value refers to the degree of performance exhibited by a model following each iteration of optimization. As seen in Fig. 4 Loss of a CNN model., the training loss in the CNN model was 0.0407, while the validation loss was 0.0384.

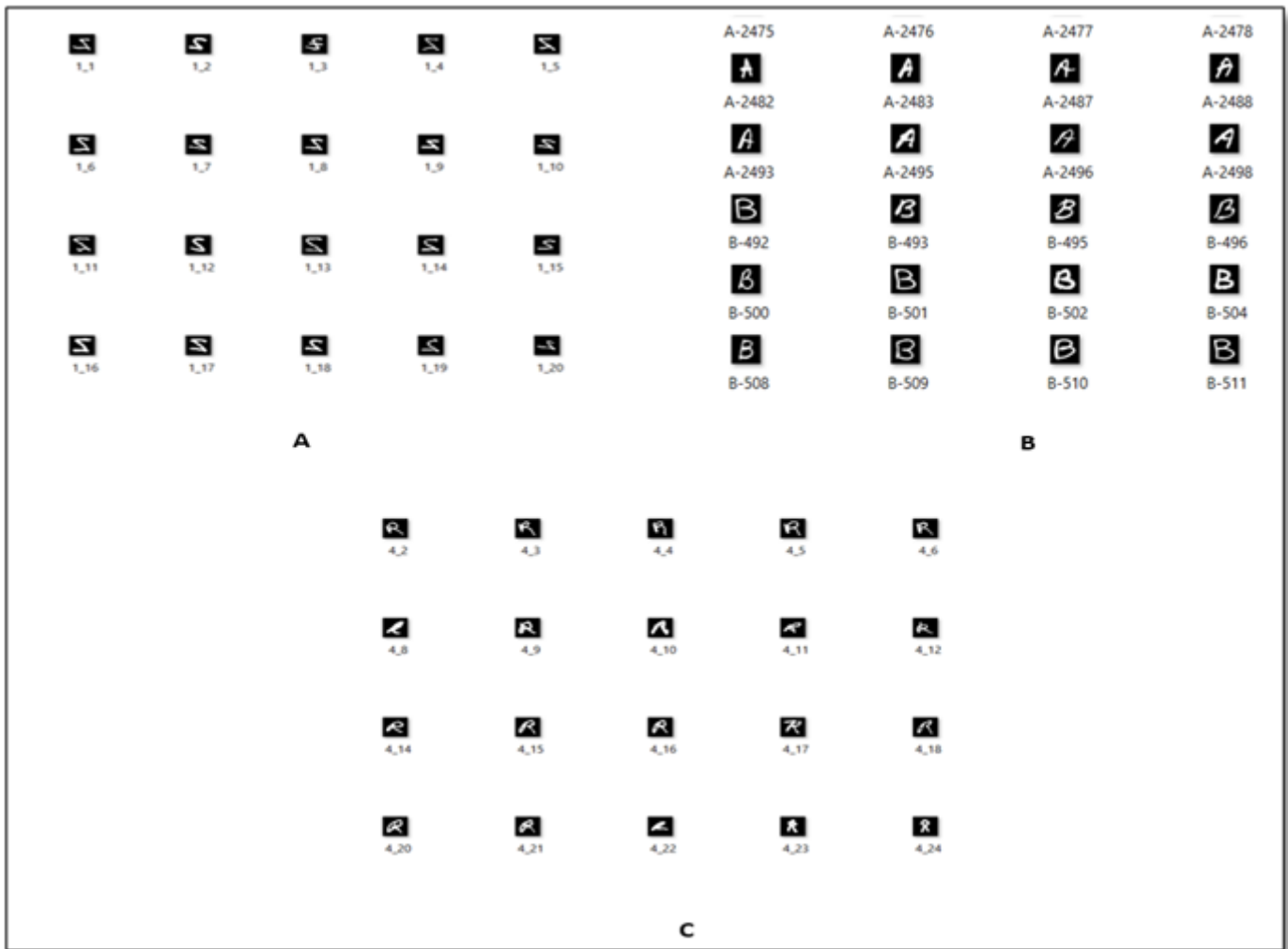


Fig. 2. Samples of datasets used.

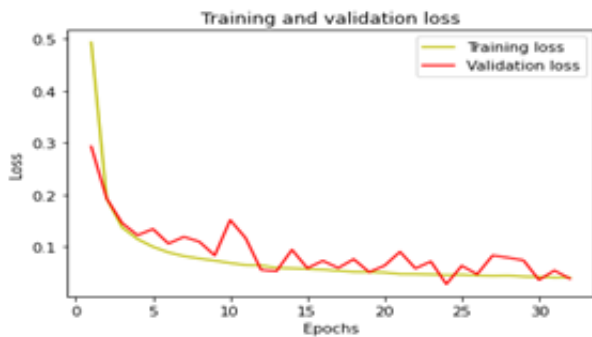


Fig. 4. Loss of a CNN model.



Fig. 5. Accuracy of CNN model.

The evaluation of a model’s accuracy often occurs subsequent to the estimation of its parameters and is quantified as a percentage. The accuracy of a model’s predictions is determined by its ability to closely align with the actual data. Fig. 5 Accuracy of CNN model. clarifies the accuracy of the CNN model, which reached above 98.59% in feature extraction with batch size = 64 and number of epochs = 40.

Confusion matrix: A confusion matrix is a tabular representation that provides a concise summary of the predictive performance of an ML model when evaluated against a specific set of test data. Measurement of classification model performance

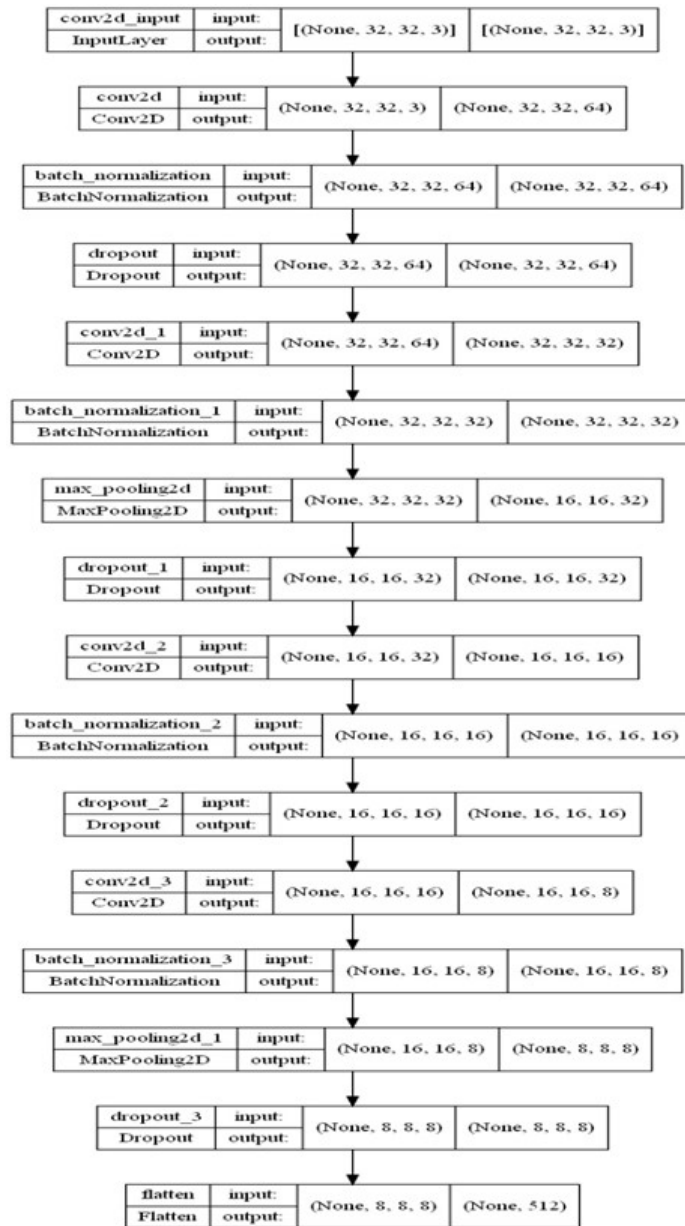


Fig. 3. CNN layers.

is frequently employed to assess the accuracy of predicting categorical labels for input instances. This provides a means of comparing the observed values with the values that were forecasted or estimated. The confusion matrix of the CNN model is shown in Fig. 6CNN confusion matrix. and 7CNN report of confusion matrix., where 0 denotes normal, 1 point denotes corrected, and 2 points denote reversal. This matrix is particularly useful for measuring important metrics: recall, precision, specificity, and accuracy. The elements within the matrix are classified as true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). A TP occurs when an observation is classified as positive and is correctly expected to be positive. An FN occurs when an observation is determined to be positive but is incorrectly anticipated to be negative. A TN occurs when the observed outcome is negative and is accurately anticipated to be negative. An FP

occurs when an observation is determined to be negative but is incorrectly anticipated to be positive.

$$\begin{aligned}
 \text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN} \\
 \text{Precision} &= \frac{TP}{TP + FP} \\
 \text{Recall} &= \frac{TP}{TP + FN} \\
 \text{F1-Score} &= \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{Recall}}
 \end{aligned}
 \tag{2}$$

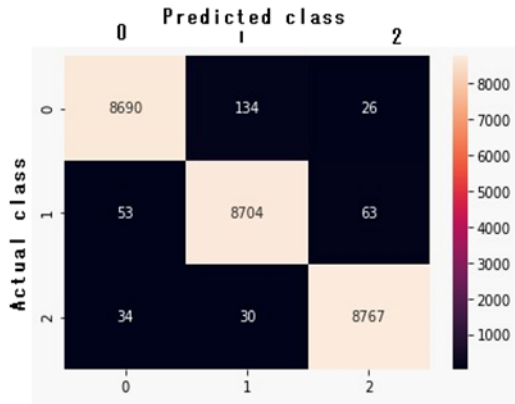


Fig. 6. CNN confusion matrix.

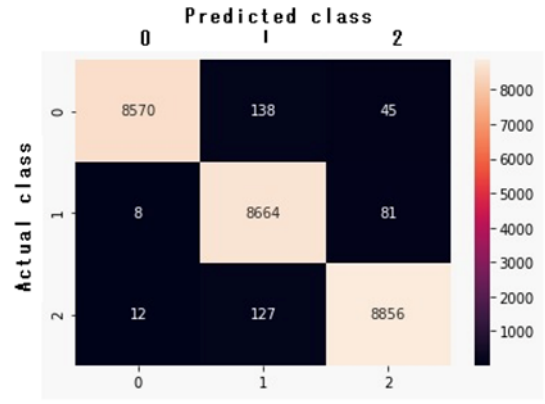


Fig. 9. CNN-RF confusion matrix.

The CNN-SVM model achieved high accuracy—99.33%—as illustrated in Fig. 10 CNN-SVM report of confusion matrix. and 11 CNN-SVM confusion matrix.. This proves what noted in [21]–[24] that SVM classifiers have a strong generalization ability.

	precision	recall	f1-score	support
Corrected	0.99	0.98	0.99	8850
Normal	0.98	0.99	0.98	8820
Reversal	0.99	0.99	0.99	8831
accuracy			0.99	26501
macro avg	0.99	0.99	0.99	26501
weighted avg	0.99	0.99	0.99	26501

Fig. 7. CNN report of confusion matrix.

	precision	recall	f1-score	support
Corrected	1.00	0.99	1.00	8753
Normal	0.99	0.99	0.99	8753
Reversal	0.99	0.99	0.99	8995
accuracy			0.99	26501
macro avg	0.99	0.99	0.99	26501
weighted avg	0.99	0.99	0.99	26501

Fig. 10. CNN-SVM report of confusion matrix.

In the CNN-RF model, the accuracy as we see in Fig. 8 CNN-RF report of confusion matrix. and Fig. 9 CNN-RF confusion matrix. achieved 98.44%

	precision	recall	f1-score	support
Corrected	1.00	0.98	0.99	8753
Normal	0.97	0.99	0.98	8753
Reversal	0.99	0.98	0.99	8995
accuracy			0.98	26501
macro avg	0.98	0.98	0.98	26501
weighted avg	0.98	0.98	0.98	26501

Fig. 8. CNN-RF report of confusion matrix.

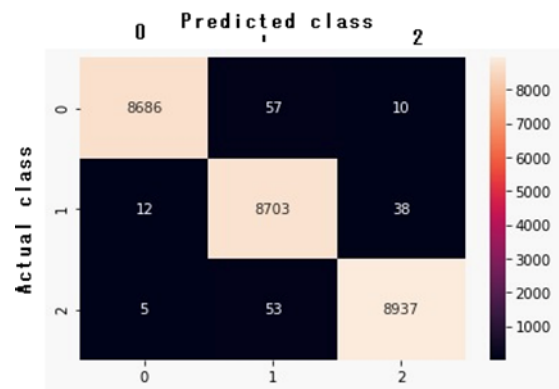


Fig. 11. CNN-SVM confusion matrix.

Comparison: Comparison of the Proposed models with the models in previous studies on the same dataset demonstrated in Table II Comparison of the Proposed Models with Previous Studies Models which demonstrate the superiority of our study models.

TABLE II. COMPARISON OF THE PROPOSED MODELS WITH PREVIOUS STUDIES MODELS

Models	Accuracy
LeNet-5	0.8873
Modified LeNet-5	0.9534
CNN	0.9859
CNN-RF	0.9844
CNN-SVM	0.9933

VII. CONCLUSION

In the field of image identification, the DL architecture known as CNN is becoming remarkably more significant. It has been utilized in all previous studies to recognize dyslexia through handwriting. Its performance has varied, starting at 55% and going up to 95% accuracy. This has been due to factors such as the small-scale size of the data set as well as problems with imbalanced classes. Our paper targeted developing the CNN model to maximize the performance in feature extraction and therefore classification and leveraging the combination of the DL and ML models to improve the prediction of dyslexia through handwriting image models, in terms of loss and accuracy, in the training model test. CNN-SVM outperformed CNN and CNN-RF, which reached 98.59% and 98.44%, respectively, while CNN-SVM achieved 99.33% in multiclass classification. Expanding research and the development of applications based on the identification of dyslexia through online handwriting is possible with the development of DL. We encourage researchers to seek to build a handwriting data set for children with dyslexia, as there is no collection currently available, except for those to whom the research has been applied.

REFERENCES

[1] 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," in *2021 11th IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE 2021*, 2021, pp. 194–199.

[2] Y. S. Kim, R. K. Wagner, and D. Lopez, "Developmental relations between reading fluency and reading comprehension: A longitudinal study from grade 1 to grade 2," *Journal of Experimental Child Psychology*, vol. 113, no. 1, pp. 93–111, 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.jecp.2012.03.002>

[3] K. Spoon, D. Crandall, and K. Siek, "Towards detecting dyslexia in children's handwriting using neural networks," in *Int. Conf. Mach. Learn. AI Soc. Good Work.*, 2019, pp. 1–5.

[4] G. H. Kwak, C. W. Park, K. D. Lee, S. I. Na, H. Y. Ahn, and N. W. Park, "Potential of hybrid cnn-rf model for early crop mapping with limited input data," *Remote Sensing*, vol. 13, no. 9, 2021.

[5] N. D. Alqahtani, B. Alzahrani, and M. S. Ramzan, "Deep learning applications for dyslexia prediction," *Appl. Sci.*, vol. 13, no. 5, pp. 1–17, 2023.

[6] K. Spoon, K. Siek, D. Crandall, and M. Fillmore, "Can we (and should we) use ai to detect dyslexia in children's handwriting?" in *Artif. Intell. Soc. Good (NeurIPS 2019)*, 2019, pp. 1–6.

[7] I. S. Isa, W. N. S. Rahimi, S. A. Ramlan, and S. N. Sulaiman, "Automated detection of dyslexia symptom based on handwriting image for primary school children," *Procedia Comput. Sci.*, vol. 163, pp. 440–449, 2019.

[8] P. Yogarajah and B. Bhushan, "Deep learning approach to automated detection of dyslexia-dysgraphia," in *The 25th IEEE International Conference on Pattern Recognition*, 2020, pp. 1–12. [Online]. Available: <https://www.micc.unifi.it/icpr2020/>

[9] I. S. Isa, M. A. Zahir, S. A. Ramlan, W. Li-Chih, and S. N. Sulaiman, "Cnn comparisons models on dyslexia handwriting classification," *ES-TEEM Acad. J.*, vol. 17, no. March, pp. 12–25, 2021.

[10] H. Derrouz, A. Cabri, H. Ait Abdelali, R. Oulad Haj Thami, F. Bourzeix, S. Rovetta, and F. Masulli, "End-to-end quantum-inspired method for vehicle classification based on video stream," *Neural Computing and Applications*, vol. 34, no. 7, pp. 5561–5576, 2022.

[11] N. Altwaijry and I. Al-Turaiki, "Arabic handwriting recognition system using convolutional neural network," *Neural Comput. Appl.*, vol. 33, no. 7, pp. 2249–2261, 2021.

[12] W. Yang, D. Dai, B. Triggs, and G. S. Xia, "Sar-based terrain classification using weakly supervised hierarchical markov aspect models," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4232–4243, 2012.

[13] E. Othman, Y. Bazi, N. Alajlan, H. Alhichri, and F. Melgani, "Using convolutional features and a sparse autoencoder for land-use scene classification," *Int. J. Remote Sens.*, vol. 37, no. 10, pp. 2149–2167, 2016.

[14] C. K. Dewa, A. L. Fadhilah, and A. Afiahayati, "Convolutional neural networks for handwritten javanese character recognition," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 12, no. 1, p. 83, 2018.

[15] S. Joseph and J. George, "Handwritten character recognition of modi script using convolutional neural network based feature extraction method and support vector machine classifier," in *2020 IEEE 5th Int. Conf. Signal Image Process. ICSIP 2020*, 2020, pp. 32–36.

[16] L. Chen, S. Wang, W. Fan, J. Sun, and N. Satoshi, "Reconstruction combined training for convolutional neural networks on character recognition," in *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 2015-Novem, 2015, pp. 431–435.

[17] D. Decoste and B. Schölkopf, "Training invariant support vector machines," *Mach. Learn.*, vol. 46, no. 1–3, pp. 161–190, 2002.

[18] M. Z. Tsegaye and P. M. Shashi, "A new hybrid cnn-svm model for amharic character image recognition," vol. 9, no. 10, pp. 397–413, 2021.

[19] M. Elleuch, N. Tagougui, and M. Kherallah, "A novel architecture of cnn based on svm classifier for recognising arabic handwritten script," *Int. J. Intell. Syst. Technol. Appl.*, vol. 15, no. 4, pp. 323–340, 2016.

[20] A. A. A. Ali and S. Mallaiah, "Intelligent handwritten recognition using hybrid cnn architectures based-svm classifier with dropout," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 6, pp. 3294–3300, 2022.

[21] H. Wu, D. Li, and M. Cheng, "Chinese text classification based on character-level cnn and svm," *Int. J. Intell. Inf. Database Syst.*, vol. 12, no. 3, pp. 212–228, 2019.

[22] S. S. Rosyda and M. I. Fanany, "Handwriting recognition on form document using convolutional neural network and support vector machines (cnn-svm)," in *2017 5th Int. Conf. Inf. Commun. Technol. ICoICT 2017*, 2017.

[23] M. H. Momade, S. Shahid, M. R. bin Hainin, M. S. Nashwan, and A. Tahir Umar, "Modelling labour productivity using svm and rf: a comparative study on classifiers performance," *International Journal of Construction Management*, vol. 22, no. 10, pp. 1924–1934, 2022.

[24] V. F. Rodriguez-Galiano, B. G. B. J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 67, pp. 93–104, 2012. [Online]. Available: <https://doi.org/10.1016/j.isprsjprs.2011.11.002>

[25] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[26] S. Barton, "How to spot dyslexia in a writing sample," 2017.

[27] Dyslexia handwriting dataset. Kaggle. Accessed on Sep. 10, 2023. [Online]. Available: <https://www.kaggle.com/datasets/drizasazanitaisa/dyslexia-handwriting-dataset>

[28] Nist special database 19. NIST. Accessed on Nov. 04, 2022. [Online]. Available: <https://www.nist.gov/srd/nist-special-database-19>

[29] A-z handwritten alphabets in .csv format. Kaggle. Accessed on Nov. 04, 2022. [Online]. Available: <https://www.kaggle.com/datasets/sachinpatel21/az-handwritten-alphabets-in-csv-format>

[30] M. A. H. Akhand, M. Ahmed, M. M. H. Rahman, and M. M. Islam, "Convolutional neural network training incorporating rotation-based generated patterns and handwritten numeral recognition of major indian scripts," *IJETE J. Res.*, vol. 64, no. 2, pp. 176–194, 2018.

- [31] S. S. Rosyda and T. W. Purboyo, "A review of various handwriting recognition methods," *Int. J. Appl. Eng. Res.*, vol. 13, no. 2, pp. 1155–1164, 2018.
- [32] J. Koushik, "Understanding convolutional neural networks using a minimal model for handwritten digit recognition," arXiv preprint, 2016. [Online]. Available: <http://arxiv.org/abs/1605.09081>
- [33] J. Bjorck, C. Gomes, B. Selman, and K. Q. Weinberger, "Understanding batch normalization," *arXiv preprint*, vol. 4, no. NeurIPS, pp. 1–24, 2018.
- [34] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognition Letters*, vol. 141, pp. 61–67, 2021.
- [35] J. Bjorck, C. Gomes, B. Selman, and K. Q. Weinberger, "Understanding batch normalization," *arXiv preprint*, vol. 4, no. NeurIPS, pp. 1–24, 2018.
- [36] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognition Letters*, vol. 141, pp. 61–67, 2021.