Hindawi
Journal of Sensors
Volume 2023, Article ID 6897719, 13 pages
https://doi.org/10.1155/2023/6897719

Research Article

Tamil OCR Conversion from Digital Writing Pad Recognition Accuracy Improves through Modified Deep Learning Architectures

V. Jayanthi and S. Thenmalar

Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India

Correspondence should be addressed to V. Jayanthi; jv4352@srmist.edu.in

Received 16 May 2023; Revised 3 July 2023; Accepted 5 July 2023; Published 4 August 2023

Academic Editor: Huan Liu

Digital handwritten recognition is an emerging field in optical character recognition (OCR). A digital writing pad replaces manual writing. In digital writing, the alphabet changes in font and shape. During OCR recognition, covert text file errors occur due to digital pen pressure and digital pen position on the digital pad by the writer. The shape changes in the alphabet lead to an error during the conversion of OCR to text. The above problem arises in Tamil, Chinese, Arabic, and Telugu, where the alphabet consists of bends, curves, and rings. OCR-to-text conversion for the Tamil language has more word errors due to angles and curves in the alphabet, which need to be converted accurately. This paper proposes ResNet two-stage bottleneck architecture (RTSBA) for Tamil language-based text recognition written on a digital writing pad. In the proposed RTSBA, two separate stages of neural networks reduce the complexity of the Tamil alphabet recognition problem. In the initial stage, the number of inputs and variables is reduced. In the final stage, time and computation complexity are reduced. The proposed algorithm has been compared with traditional algorithms such as long short-term memory, Inception-v3, recurrent neural networks, convolutional neural networks, and a two-channel and two-stream transformer. Proposed methods, such as RTSBA applied in the digital writing pad-handwritten and HP lab datasets, achieved an accuracy of 98.7% and 97.1%, respectively.

1. Introduction

Tamil is spoken in different countries, such as southern India and worldwide [1]. Tamil is spoken and written in Tamil Nadu, India, and other countries such as Singapore, Sri Lanka, and Malaysia [2]. Tamil literature is from 300 BC to now and is an ancient language. From 700 to 1,600 AD, the Tamil language was called as Middle Tamil. In the year 1600, the Tamil language was identified as modern Tamil. The old Tamil writings are carved in stones; the middle Tamil was written on palm leaves. Modern Tamil writing is in the textbook, which is written using pen and paper. Recently, Tamil has been written on a digital writing pad, and Tamil documentation is stored in digital formats. A digital writing pad saves paper, and documents are easily accessible and cost-effective.

Handwriting recognition in the Tamil language is an ongoing challenge. Researchers developed algorithms for Tamil handwriting recognition using convolutional neural networks (CNNs). Accuracy in recognition needs to be increased. Tamil is a script-based language that is written in different styles, and recognizing handwritten Tamil text accurately becomes a challenge. Researchers improve handwriting recognition through different methods, such as font normalization, character segmentation, and model ensembling.

Tamil handwriting recognition is a challenging task due to the presence of complex characters and patterns. Studies have shown that deep learning-based methods such as CNNs and recurrent neural networks (RNNs) can achieve high-accuracy outcomes for Tamil handwriting recognition. Moreover, feature extraction and image segmentation extract characters from handwritten text images and improve handwriting recognition accuracy [3]. Other methods, such as bidirectional

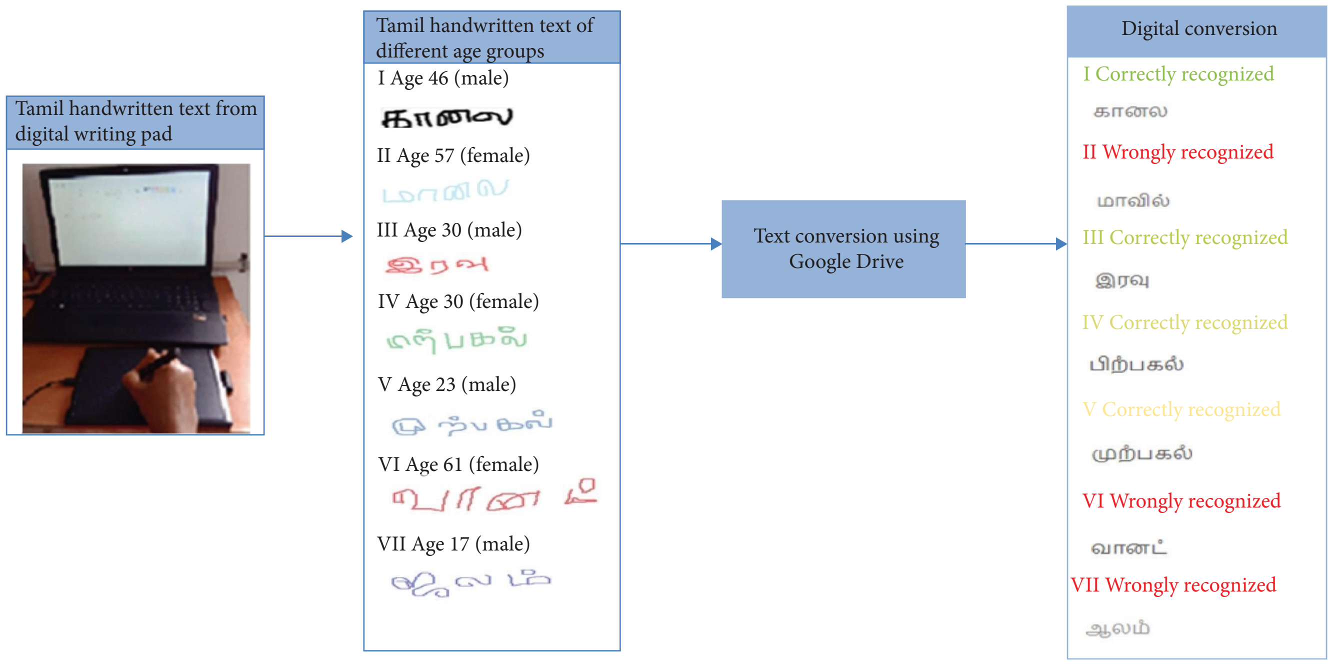


FIGURE 1: Tamil handwritten text from different age groups from a digital writing pad to conversion using Google Docs.

long short-term memory (BiLSTM) encoding, are used for Tamil handwriting with high recognition accuracy.

Automatic character recognition and conversion of online handwritten text compression with optical character recognition (OCR) for the Tamil alphabet is challenging. Writing on the tablet using the tip of your fingers or a stylus pen has increased due to the growth of internet technology. A novel method is required for Tamil handwriting text segmentation and classification. The major problems with Tamil handwriting text are size variations, dimensional changes, irregular stylus points, discontinuity of structures, superfluous overloops, shape variation, and distinctive curves [4]. Figure 1 shows the Tamil handwritten to digital writing pad for different age groups and the optical digital conversion system. A handwritten document with a digital pen and pad shows the wrong recognition of the word. To overcome the above problem, a ResNet two-stage bottleneck architecture (RTSBA) was proposed. The two-stage bottleneck architecture of ResNet is designed for filtering the noise and enhancing the image for better network performance.

In addition, nonlocal (NL) and attention-gate (AG) blocks enhance performance by increasing the network's ability to capture long-range dependencies. The bottleneck structure of ResNet reduces the number of parameters in a network and increases the accuracy. NL and AG blocks reduce the parameter's computational complexities and memory. A two-stage bottleneck architecture has fewer parameters and leads to fast computation. The reduced number of parameters in the bottleneck architecture has reduced overfitting.

Contributions are: (1) To recognize handwritten text in the digital pad and pen text recognition using RTSBA, which is based on segmentation. (2) To recognize the handwritten text using the digital pad and pen at different pen pressures using the proposed RTSBA algorithm, which segments the alphabet into simple curves, and closed simple curves for recognition. (3) To recognize digital writings in the Tamil language based on demography, such as in Madurai, Tirunelveli, Coimbatore, and Thanjavur in Tamilnadu, where the alphabet changes based on the non-simple curve in the Tamil alphabet. (4) To recognize digitized writings in the Tamil language through syllable-based alphabet detection and classification with RTSBA. (5) To classify and recognize the Tamil language alphabet based on the different age groups of writing, such as (i) 15–25, (ii) 26–35, (iii) 36–45, (iv) 46–55, and (v) 56–65, and compare the Accuracy, Precision, Recall, and F1 score of the proposed method with traditional algorithms.

2. Related Work

A novel deep learning-based approach for multilingual handwritten numeral recognition has been developed [5]. The process involves a pretrained CNN model and transfer learning. The developed method is tested on eight images collected from eight different languages: Arabic, English, Persian, Hebrew, Urdu, Mongolian, Kalmyk, and Spanish. Moreover, the algorithm provides near-perfect results in specific languages, such as Arabic and Persian. The current study never included any handwriting variability analysis and evaluated the network capacity to recognize handwritten digits of people from different countries or writing styles. A fully convolutional recurrent network (FCRN) [6] recognizes online handwritten Chinese text by incorporating the spatial-semantic context. FCRN is trained from end-to-end with multistage training using deep learning techniques for feature extraction and mapping the data to a higher semantic space. The model is prone to vanishing gradients, limiting performance on more complex tasks. An offline handwritten Chinese text recognition technique based on a fully CNN is

Table 1: Comparison of different methods and languages.

References	Model	Language	Dataset	Type	Dataset size	Writing devices
[16]	CNN	Arabic	CMATERDB	Digits	3,000	Pen and Paper
[17]	CNN BiLSTM	Arabic	KHATT	Word	325 paragraph	Pen and Paper
[18]	CNN-TL	Bangla and Devanagiri	CMATERDB ICDAR Ph.D. Indic_11	Word	361 handwritten documents	Pen and Paper
[19]	GAN	Chinese	HIT-MW and ICDAR	Word	1,003 handwritten images	Pen and Paper
[20]	CNN	Telugu	Own dataset	Character	16 characters and 21 guninthalu, total of 275,520	Pen and Paper Touchpad device
Proposed	RTSBA	Tamil	DWP-H and HP Lab datasets	Word	251 writers total of 85,800 images	Tablet

CNN, convolutional neural network; BiLSTM, bidirectional long short-term memory; RTSBA, ResNet two-stage bottleneck architecture; DWP-H, digital writing pad-handwritten; GAN, generative adversarial network; CMATERDB, center for microprocessor applications for training education and research data base; KHATT, KFUPM Handwritten Arabic TexT; ICDAR, international conference on document analysis and recognition; HIT-MW, harbin institute of technology-multiple writers.

suggested [7]. This technique uses CNNs and classifies handwritten Chinese text input in images. The model consists of several convolution layers, residual blocks, and attention mechanisms. The convolution layers are used to extract features from images, while the residual blocks increase the information capacity of the model. Finally, the attention mechanism allows the model to focus on the most critical information in the image and never recognize the sequence of characters in a sentence, recognizing only the individual characters.

BiLSTM with data augmentation, including rotating, shifting, and stretching, improves text recognition accuracy, and the model focuses on a specific type of handwriting [8]. Deep feature learning on wearable sensors improves handwritten character recognition [9]. The model relies on a deep learning architecture and extracts a set of high-level feature representations, which are used to classify handwritten characters. The features are utilized in a supervised learning approach to predict the handwritten character. The model needs to be more balanced, as the timing and sequencing information is complex and maps accurately to handwriting features.

The bottleneck transformer (BNT) [10] is a novel network architecture developed for visual recognition tasks. It uses a simplified transformer-based encoder–decoder network and compresses the image features at multiple scales before feeding them into a cross-scale self-attention module for efficient inference. BNT uses a block sparsity regularization technique and reduces the complexity of the network. Furthermore, the faster deduction is performed. The compression enables BNT for accurate classification in unseen images with a much smaller number of parameters than the standard Transformer-based. It never considers relative positional information, which limits the performance of tasks with highly structured data. A combination of CNNs and RNNs [11] provides the best accuracy in classifying and recognizing handwritten mathematical symbols and expressions, but is limited to a few symbols and expressions and never performed for mathematical operations. An RNN-based deep learning model is used for the accurate classification of text and non-text strokes in online handwritten Devanagari documents [12]. The model is trained using multiple two-dimensional feature vectors extracted from the stroke images and evaluated for average accuracy, obtained by selecting the model's best-performing parameters. The model classifies the strokes as text or non-text.

A novel approach recognizes handwritten offline Tamil characters using the combination of conditional generative adversarial network (cGAN) and CNN [13]. The preprocessing step involves identifying the characters from the handwritten dataset using edge detection methods. Using cGAN, the dataset is augmented and generates more primitive features used to train the CNN model. Finally, the model is tested on the augmented images and evaluated for performance and accuracy. The method is unable to support complex writing.

To improve the efficiency of online handwriting recognition (OHR), Carbune et al. [14] suggested "Fast Multilanguage LSTM-based Online Handwriting Recognition." This approach utilizes an optimized version of the LSTM network, which runs on multiple platforms. The OHR system processes the data more rapidly than traditional LSTM networks. The model performs better for datasets with high-quality images, and noise in images affects the performance and accuracy. In [15], a segmentation method based on an attention-embedded lightweight network is developed. The network is a combination of CNNs and attention modules and captures more informative features from images and improves segmentation accuracy, and the accuracy of the segmentation varies based on the types of features in training the model. Table 1 shows the different methods and languages.

3. RTSBA Method for Non-Simple and Straightforward Curve Text Classification

In this paper, the proposed RTSBA model is shown in Figure 2, which is used to recognize Tamil text; text images are collected from different sources, such as offline note-taking and online sources, and text is written on paper and with a pen. RTSBA segmentation consists of two separate stages, reducing Tamil handwriting recognition's complexity. The two-stage bottleneck architecture segments the text more accurately because the network focuses on the important features and ignores the irrelevant objects. In addition, accurate and detailed segmentation results are compared with single-

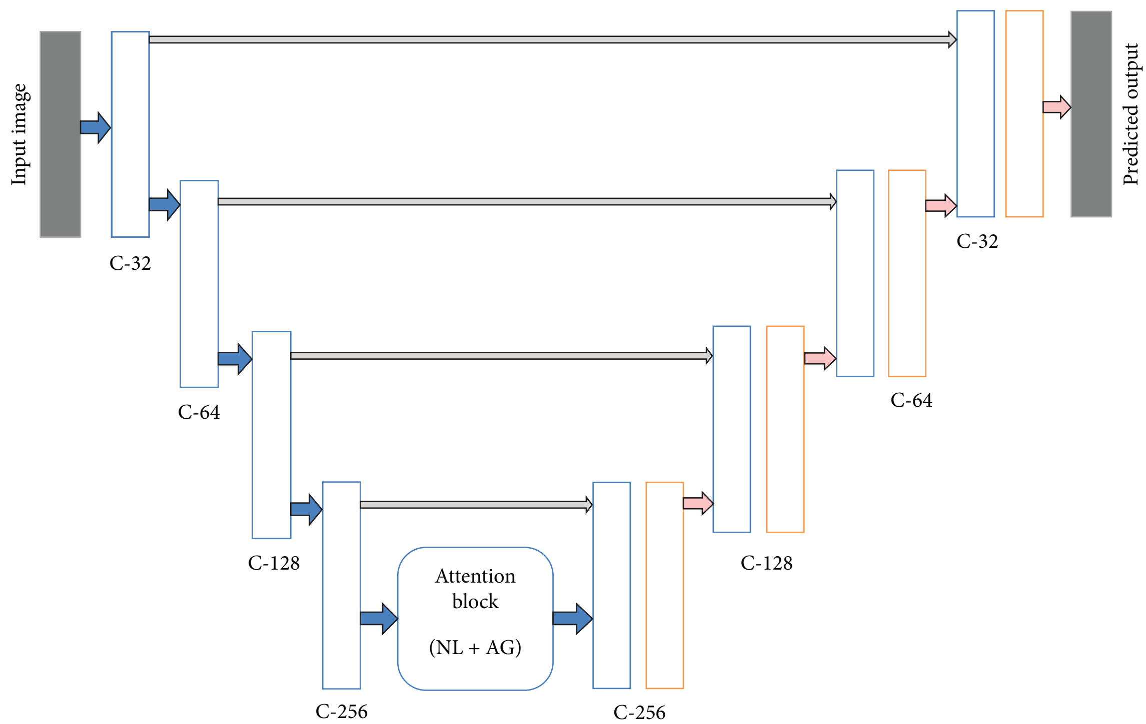


FIGURE 2: ResNet two-stage bottleneck architecture (RTSBA).

stage architectures. RTSBA reduces complexity through different types of neural networks at each stage and more accurately predicts curves in the alphabet. RTSBA consists of an encoder and decoder phase with an attention block between the steps. The attention block is integrated with the NL and AG blocks, which overcome the bottleneck problems in CNN.

The encoder phase consists of convolution layers and max pool layers, which acquire the content of the images. Convolution layers capture the image features, followed by the RELU and max pools to dilute feature parameters. Squeeze and excitation blocks are introduced in the encoder phase to overcome feature loss and workload problems. After two convolution layers, the image enters the network; two-channel separation is used. A down-sampling process makes the image the same size as the input image. The model learns the detailed image properties, and the image size increases using the upsampling process.

The feature map is adjusted to a size similar to the decoder phase input image. The decoder phase consists of four upsampling blocks. Each sampling block contains two convolution layers and one RELU layer. The size of input information decreases in the encoder phase and increases in the decoder phase. Some attributes are transmitted if the input is compressed and a bottleneck occurs. To overcome the bottleneck problems, a two-stage block is developed to minimize loss of input.

In Figure 3, the NL block feature map is represented; $\otimes$ and $\oplus$ depict multiplication and addition, respectively. For each row, the softmax operation is performed block by block as described in Equation (1).

$$C_i = W_z b_i + a_i. \tag{1}$$

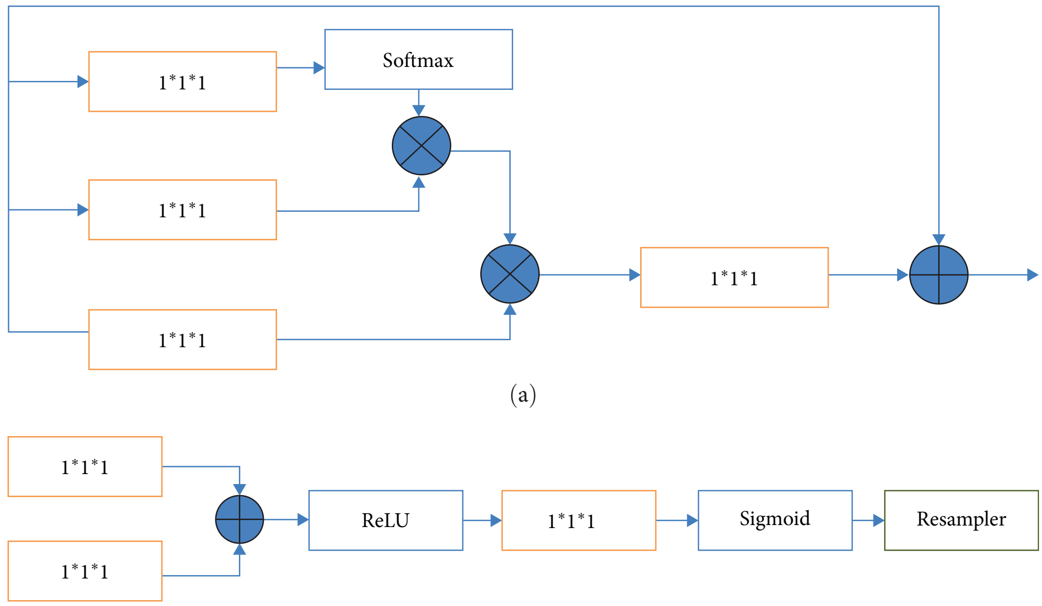
W_z represents the initial weight values; a_i is for residual information; b_i is for similar size information; and c_i is for block value. The bottleneck problem is solved using the NL and AG blocks.

The AG block has 1×1×1 convolution. The convolution is resampled with the RELU activation function through the sigmoid function. NL blocks help the networks by capturing the long-range dependencies in their data. Long-range dependencies refer to relationships between data points that are far apart. These relationships play a vital role in image classification. This relationship is never used in traditional convolutional networks. NL blocks capture dependencies by utilizing self-attention.

NL blocks improve computational efficiency. Since self-attention modules never rely on convolutional kernels or local processing, they perform better with fewer operations. NL blocks are more powerful and accurate than traditional convolutional layers. The self-attention module allows for global context-aware learning and will enable networks to capture relationships between different locations in the image. This led to more accurate predictions and better performance. The AG block helps focus on the input's relevant features and removes irrelevant distractions, improving the model's performance. AGs reduce the computational cost through faster convergence and improve the model's performance. The AG block provides a visual representation of features and makes it easier to interpret the model to debug potential problems.

4. Experiment and Result

The digital writing pad-handwritten (DWP-H) dataset is created using the Tablet Model Number (Wacom CTL-672/K0-CX), a graphic tablet for online and offline. Text datasets were acquired during ambient lighting conditions and pressure-

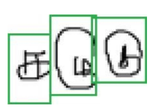
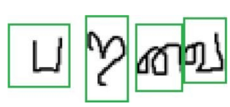
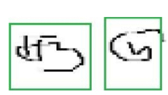


(a)

(b)

Figure 3: Two-stage bottleneck structure (a) nonlocal block, (b) attention-gate block.

Table 2: Result comparison CNN, MMU-SNet, and RTSBA.

Tamil handwritten word	Segmented	Correct word	CNN	MMU-SNet	RTSBA
		கழுகு	கமுகு ✗	கழுகு ✓	கழுகு ✓
		பறவை	பநவை ✗	பரவை ✗	பறவை ✓
		சூடு	ஆடு ✗	ஞடு ✗	சூடு ✓

CNN, convolutional neural network; RTSBA, ResNet two-stage bottleneck architecture; MMU-SNet, modified multi-scale segmentation network.

based stylus writing on a digital pad. A total of 251 writers' samples were obtained in *.tiff image format. Different age groups based on text images are collected, such as children aged between 10 and 18, adults aged between 19 and 59, and old males and females aged between 60 and 75. Samples per class were 5, few contributed as many as ten, and the number of samples per class was 550 with a character size of 92×133. The images were resized; the longer side length is about 50×50 pixels. These images normalize by transforming each gray-scale pixel value from the $[0, 1]$ range to the $[-1, 1]$ range. The experiments used a learning rate of 0.0001 and an Adam optimizer. The network has trained for 64 batch sizes and 50 epochs. The proposed and RTSBA are analyzed using "Accuracy," "Precision," "Recall," and "F1 score" for Tamil handwritten words. The Tamil language is composed of more similar characters as well as strokes such as right curve, left curve, circle, up the curve, down the curve, dot, question mark, slanting line, standing line, sleeping and standing line, springs, down curve and circle, standing line, sleeping line. Among the above strokes, the recognition of curves is complex due to rings in the alphabet. The CNN approach is used for Tamil handwritten word recognition. Table 2 represents the proposed RTSBA results compared with modified multi-scale segmentation network (MMU-SNet) and CNN. The first-word second character is incorrectly predicted as "மு" instead of being "ழு," and the word's meaning also changed. In the second word, the second character is mispredicted as "ந" instead of "ற." The first character's incorrect prediction for the third word is "ஆ" instead of "சூ."

The performance measures are summarized for open-curve and closed-curve alphabets in the Tamil language.

5. Evaluation of the Alphabet in the Tamil Language

The Accuracy (A), Precision (P), Recall (R), and F1 score (F1) are four statistical measures to assess the performance

Table 3: Evaluation of classifier performance through statistical measures.

References	Classifiers	Accuracy	Precision (%)	Recall (%)	F1 score (%)
[21]	CNN + SALA	86.3%	88.4	94.3	91.2
[22]	AlexNet	90.5%	92.6	98.5	95.4
[22]	VGG19Net	92.5%	93.5	96.7	95.0
[23]	MMU-SNet	96.8%	99.3	97.4	98.3
Proposed	RTSBA	98.7%	99.3	99.2	99.3

MMU-SNet, modified multi-scale segmentation network.

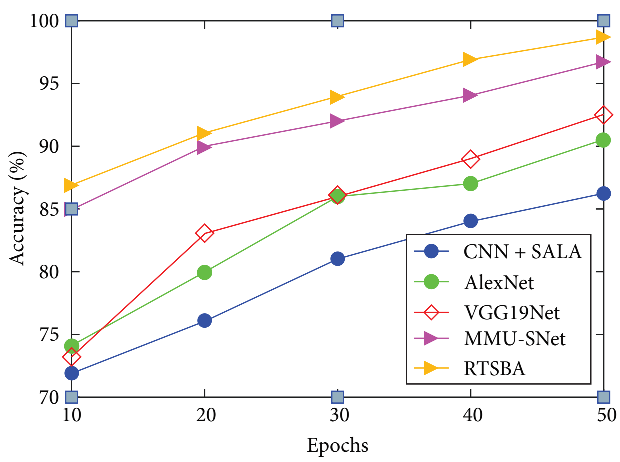


Figure 4: Evaluation of the classifier's performance concerning accuracy. MMU-SNet, modified multi-scale segmentation network.

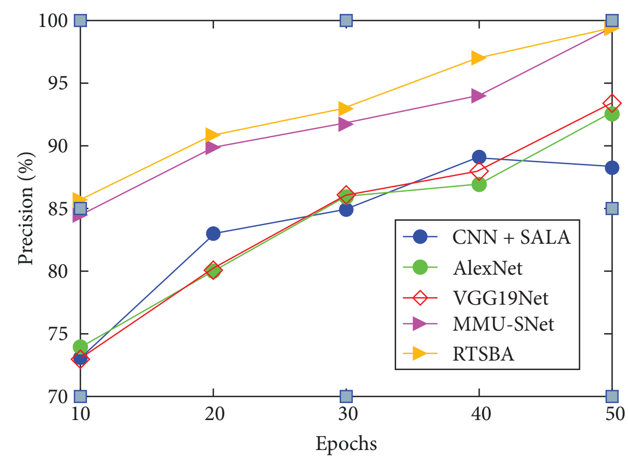


Figure 5: Evaluation of the classifier's performance concerning precision. MMU-SNet, modified multi-scale segmentation network.

of the RTSBA classifier. In Equations (2–5), the metrics are mathematically calculated. The performance evaluation of classifiers is shown in Table 3.

The model typically performs across all classes according to the Accuracy metric. Calculating accuracy involves dividing the number of correct predictions by the total number of predictions.

$$\text{Accuracy } (A) = \frac{\text{True positive (TP)} + \text{True Negative (TN)}}{\text{True positive (TP)} + \text{False Positive (FP)} + \text{True Negative (TN)} + \text{False Negative (FN)}}. \quad (2)$$

P is obtained by dividing the total number of correctly categorized positive samples by the total number of positive samples.

$$\text{Precision } (P) = \frac{\text{True positive (TP)}}{\text{True positive (TP)} + \text{False Positive (FP)}}. \quad (3)$$

Calculating the "R" is through dividing the total number of positive samples by the percentage of positive samples correctly classified as positive. The ability of a model to locate positive samples is assessed using the "R." Higher "R" values discover the positive samples.

$$\text{Recall } (R) = \frac{\text{True positive (TP)}}{\text{True positive (TP)} + \text{False Negative (FN)}}. \quad (4)$$

The weighted average of "P" and "R" is the F1. This F1 score considers both false positives and false negatives.

$$\text{F1} = 2 \times \frac{P \times R}{P + R}. \quad (5)$$

Statistical measurements are used to assess the performance of the classifiers shown in Table 3.

Figures 4–7 show the statistical measures of the proposed RTSBA methods compared with the traditional classifiers, such as the CNN, the self-adaptive lion algorithm (CNN + SALA), Vgg19Net, and AlexNet. The CNN + SALA models are computationally expensive and require more training time and resources. The performance of the CNN + SALA model is affected due to unbalanced or noisy data. Alexnet has a large number of parameters and performs less for smaller datasets. Alexnet never suits open-curve alphabet segment feature extraction. VGG19Net leads to overfitting

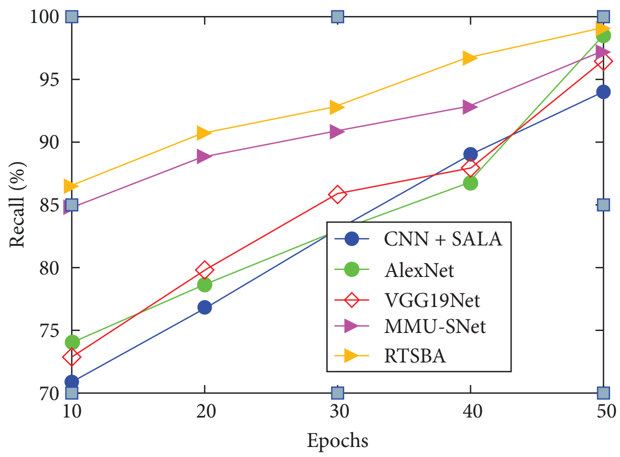


Figure 6: Evaluation of the classifier's performance with respect to recall. MMU-SNet, modified multi-scale segmentation network.

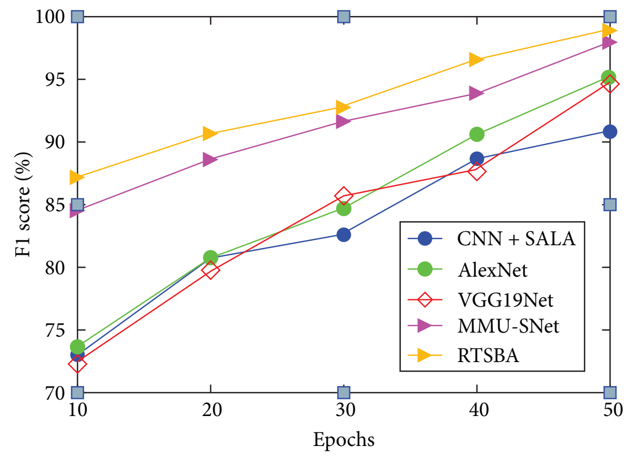


Figure 7: Evaluation of the classifier's performance concerning the F1 score. MMU-SNet, modified multi-scale segmentation network.

Table 4: Evaluation of the classifier's performance through statistical measures based on the dialect.

Dialect	Accuracy		Precision		Recall		F1 score	
	MMU-SNet (%)	RTSBA (%)	MMU-SNet (%)	RTSBA (%)	MMU-SNet (%)	RTSBA (%)	MMU-SNet (%)	RTSBA (%)
Madurai Tamil	83.9	85.9	96.1	99.2	86.2	86.3	90.9	92.3
Tirunelveli Tamil	87.1	89.7	96.3	98.5	89.6	90.7	92.8	94.5
Coimbatore Tamil	90.3	92.3	96.4	99.3	93.2	92.8	94.8	95.9
Thanjavur Tamil	93.5	95.5	98.6	99.3	94.7	96.1	96.6	97.6

RTSBA, ResNet two-stage bottleneck architecture; MMU-SNet, modified multi-scale segmentation network.

for the closed-curve alphabet. VGG19Net requires a large dataset to train.

5.1. Different Dialect Person Handwritten Tamil Alphabet Recognition. Tamil dialects are from different parts of Tamil Nadu, India, and the world. Dialects of Tamil are Madras Bashai, Kongu, Kannada, and Malayalam dialects, and writing changes from one person to another based on dialect. Each dialect has a unique vocabulary, grammar, and writing style. The Tamil diaspora speaks several Tamil dialects in other parts of the world, such as Singapore, Malaysia, and South Africa. Tamil dialects are from Madurai, Coimbatore, and Thanjavur in the Tamil Nadu state of India. The Tamil dialects vary in phonemic modifications and sound effects from Old Tami to modern Tamil. For example, the word here anku (அங்கு) is in the dialect of Coimbatore, anga (அங்க) is in the dialect of Thanjavur, and old Tamil's ankaṉa (அங்கன)—anganakula (அங்கனகுள்ள)—is a dialect of Tirunelveli, and old Tamil ankittu—the source of ankittu (அங்கிட்டு)—is a dialect of Madurai. Table 4 shows the performance analysis of the classifier based on the statistical measures of the dialect-based alphabet.

Figures 8–11 show the statistical measures of different person-written alphabet dialects, such as Madurai Tamil, Tirunelveli Tamil, Coimbatore Tamil, and Thanjavur Tamil.

5.2. Syllable-Based Alphabet Recognition. Tamil syllables are the basic units of Tamil pronunciation. They are composed

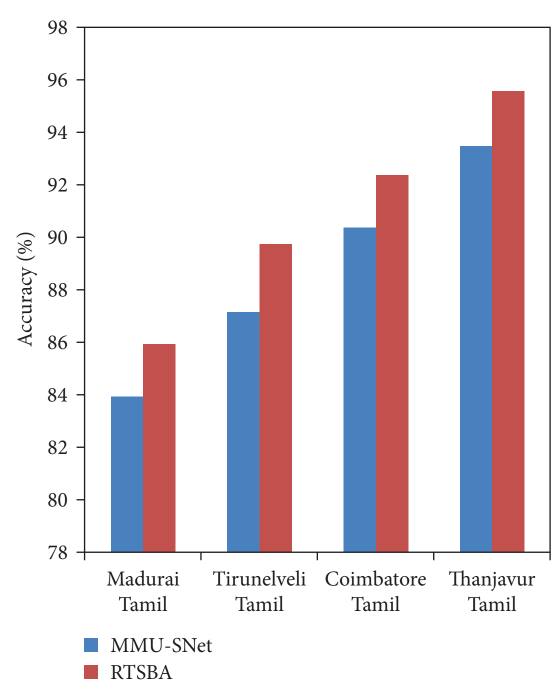


Figure 8: Evaluation of MMU-SNet and RTSBA classifier performance with different Tamil dialects and accuracy. MMU-SNet, modified multi-scale segmentation network.

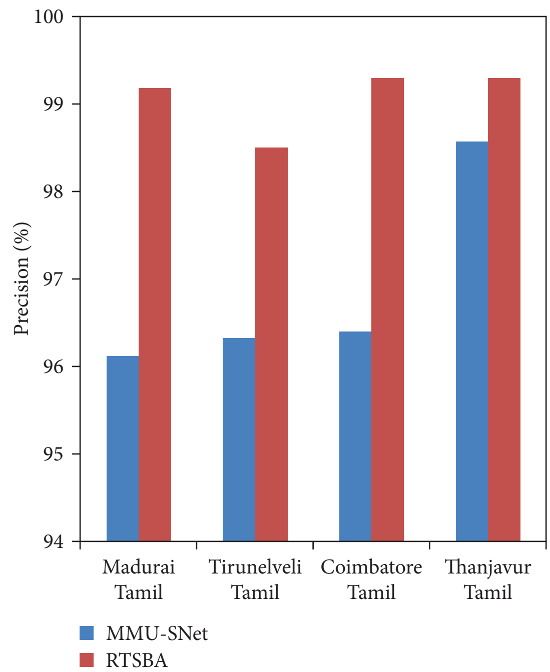


FIGURE 9: Evaluation of MMU-SNet and RTSBA classifier performance with different Tamil dialects and precision. MMU-SNet, modified multi-scale segmentation network.

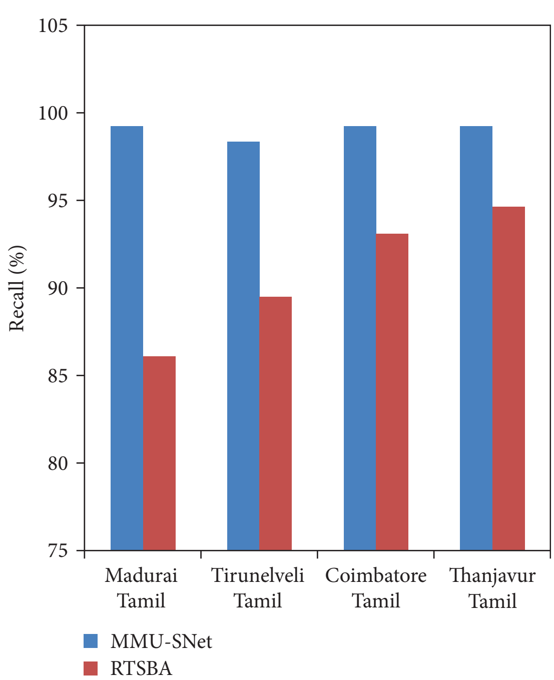


FIGURE 10: Evaluation of MMU-SNet and RTSBA classifier performance with different Tamil dialects and recall. MMU-SNet, modified multi-scale segmentation network.

of a consonant and a vowel sound, or sometimes just a vowel sound. Tamil syllables are written using a combination of twelve consonants and eighteen vowels. Tamil uses various vowel and consonant combinations to create unique characters and individual syllables. Tamil syllables are classified

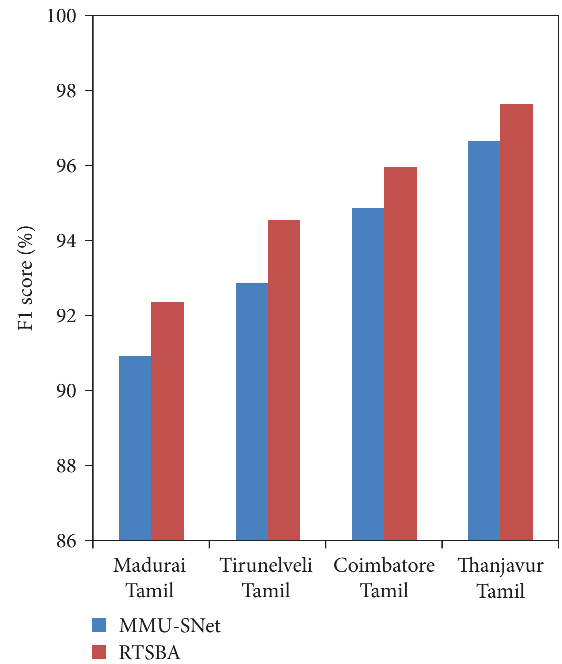


FIGURE 11: Evaluation of MMU-SNet and RTSBA classifier performance with different Tamil dialects and F1 score. MMU-SNet, modified multi-scale segmentation network.

as one, two, and three. Table 5 shows the Performance evaluation of classifiers with Accuracy, Precision, Recall, and F1 score based on a syllable.

Figures 12–15 show statistical measures of the proposed RTSBA method with classifiers such as CNN + SALA, AlexNet, and VGG19Net. The proposed RTSBA has high Accuracy, Precision, Recall, and F1 score based on syllables.

The performance evaluation of different classifiers is shown in Table 6, along with the accuracy.

Figure 16 shows the accuracy of the proposed RTSBA with DWP-H and HP labs datasets with different classifiers. The Tamil handwritten dataset from HP laboratories has 1,000 images in 169 folders, of which 550 samples are taken into account for each class, which amounts to about 156 classes. Two-stage bottleneck architectures outperform traditional architectures on tasks such as image classification and object detection. The multiple layers of the architecture provide better generalization and allow the model to perform better. The two-stage bottleneck architecture reduces the total number of parameters using smaller filters in each layer, minimizes overfitting, and makes the model more efficient. The two-stage bottleneck architecture reduces the complexity of the architecture.

5.3. Different Age-Based Handwritten Tamil Alphabet Recognition. The age group can be classified into five groups, as mentioned earlier in the section, each with different educational qualifications. Based on handwritten alphabet recognition, the performance of the classifier is assessed statistically with various age groups, as shown in Table 7.

Figures 17–20 show the statistical measures of the proposed RTSBA with different age groups based on

Table 5: Performance evaluation of classifiers through statistical measures based on syllable.

References	Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
[21]	CNN + SALA (character)	86.3	88.4	94.3	91.2
[22]	AlexNet (character)	90.5	92.6	95.5	94.3
[22]	VGG19Net (character)	92.5	93.5	96.7	95.0
[23]	MMU-SNet (syllable)	95.5	98.6	96.7	97.6
Proposed	RTSBA (syllable)	97.4	99.3	98.0	98.7

CNN + SALA, convolutional neural network, self-adaptive lion algorithm; RTSBA, ResNet two-stage bottleneck architecture; MMU-SNet, modified multi-scale segmentation network.

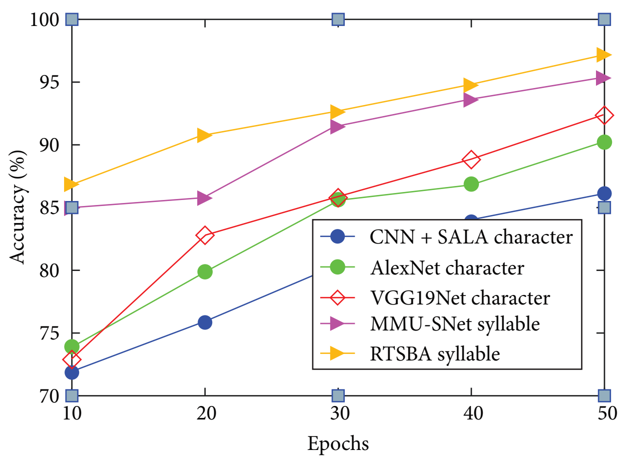


Figure 12: Evaluation of the classifier's performance and RTSBA with different methods through accuracy value. MMU-SNet, modified multi-scale segmentation network.

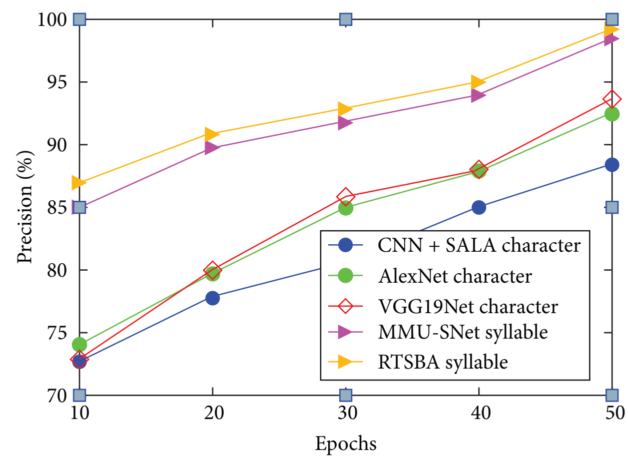


Figure 14: Evaluation of the classifiers' performance RTSBA with the different methods through recall value. MMU-SNet, modified multi-scale segmentation network.

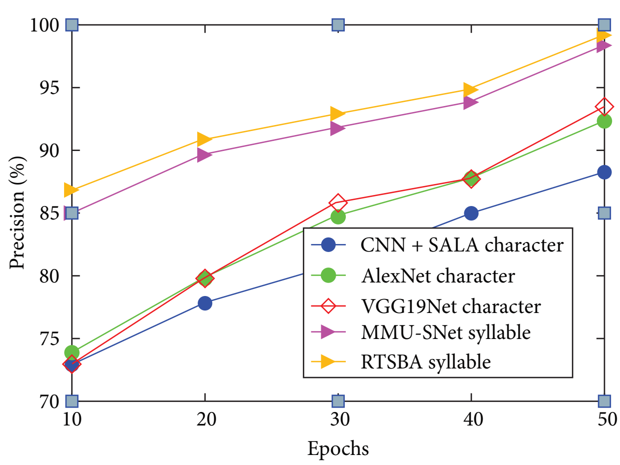


Figure 13: Evaluation of the classifier's performance and RTSBA with different methods through precision value. MMU-SNet, modified multi-scale segmentation network.

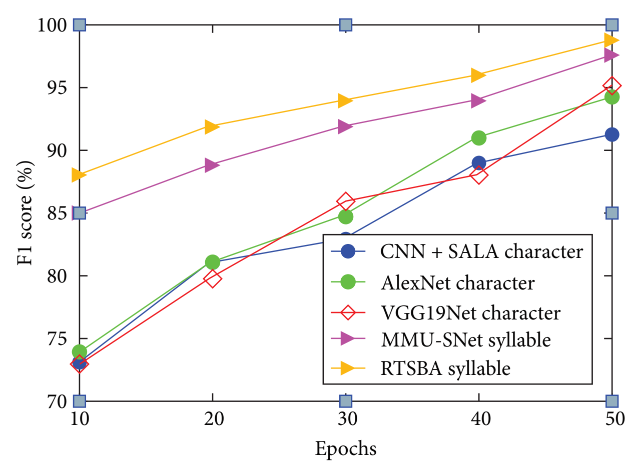


Figure 15: Evaluation of the classifiers' performance RTSBA with the different methods through the F1 score. MMU-SNet, modified multi-scale segmentation network.

handwritten Tamil alphabet recognition. The proposed RTSBA has achieved high Accuracy, Precision, Recall, and F1 score based on other age groups.

The different OCRs convert handwritten Tamil text into digital text, as shown in Table 8; these include Google Docs, i2 OCR, Easy Screen OCR, Unicode Tamil OCR, and SUBASA Tamil OCR. Figure 21 depicts OCR accuracy of 95%, 45%, 95%, 35%, and 40% for the handwritten Tamil words. The OCR, like Google Docs and Easy Screen OCR, can predict simple and complex curves, similar shapes, and discontinued curves. The OCRs such as i2 OCR, Unicode Tamil OCR, and SUBASA Tamil OCR indicate simple curves, which makes it challenging to predict complex curves, similar shapes, and discontinued curves.

Table 6: Performance evaluation of classifiers with accuracy for alphabet recognition.

References	Classifiers	Accuracy (%)
[24]	LSTM	90.3
[25]	Inception-v3	93.1
[26]	RNN	95.24
[27]	CNN	94
[28]	2C2S	93.25
Proposed	RTSBA with DWP-H dataset	98.7
	RTSBA with HP labs dataset	97.1

CNN, convolutional neural network; LSTM, long short-term memory; RTSBA, ResNet two-stage bottleneck architecture; RNN, recurrent neural network; 2C2S, two-channel and two-stream transformer; DWP-H, digital writing pad-handwritten.

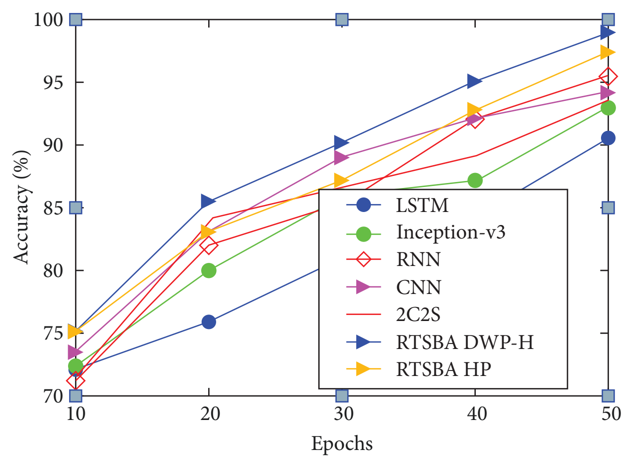


Figure 16: Evaluation of the classifier's performance with accuracy for alphabet recognition.

Table 7: Evaluation of the classifier's performance through statistical measures with different age groups based on handwritten alphabet recognition.

Age group	Accuracy		Precision		Recall		F1 Score	
	MMU-SNet (%)	RTSBA (%)	MMU-SNet	RTSBA (%)	MMU-SNet (%)	RTSBA (%)	MMU-SNet (%)	RTSBA (%)
Age 15–25	95.5	98.08	98.6%	99.35	96.7	98.7	97.6	99.02
Age 26–35	97.4	99.00	99.3%	99.04	98.0	98.7	98.7	99.35
Age 36–45	96.1	99.24	99.3	99.04	96.7	98.8	98.0	99.44
Age 46–55	96.8	99.36	98.6%	98.7	98.0	99.00	98.3	99.68
Age 56–65	94.2	97.44	97.3%	99.34	96.6	98.05	96.9	98.7

RTSBA, ResNet two-stage bottleneck architecture; MMU-SNet, modified multi-scale segmentation network.

6. Discussion

As described, the proposed RTSBA achieves a good result in predicting curves. The existing networks, such as LSTMs, are very computationally intensive and require a lot of memory to store the data in long-term memory; LSTMs are more complex than other networks and harder to tune and optimize, which has vanishing and exploding gradient problems due to the long-term dependencies of multiple update gates [24]. Inception v3 [25] is a deep learning architecture with multiple layers added to Inception v2. It is computationally expensive to train the model. Inception v3 is a black box model, making interpreting the individual layers' internal parameters impossible. The gradients vanish or explode in a long-term RNN [26]. While the gradient descent flow works for small RNNs, storing a gradient in larger RNNs is difficult, resulting in the vanishing or exploding gradient problem. This is a significant problem in RNNs. Since RNNs have many parameters, they can be prone to overfitting, affecting the network's performance and predicting new data points. For instance, context varies significantly over a period of time. Thus, RNNs never remember the exact past words or contexts, which can degrade performance. RNNs exhibit unpredictable behavior due to their internally

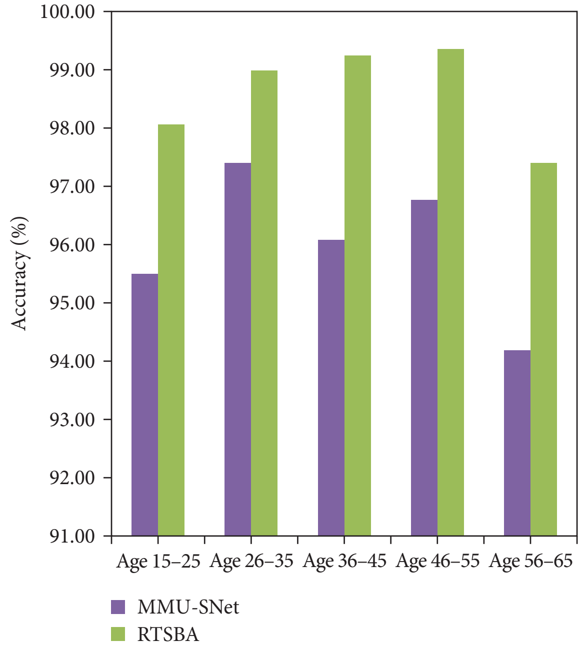


Figure 17: Evaluation of the classifier's performance with different methods through accuracy. MMU-SNet, modified multi-scale segmentation network.

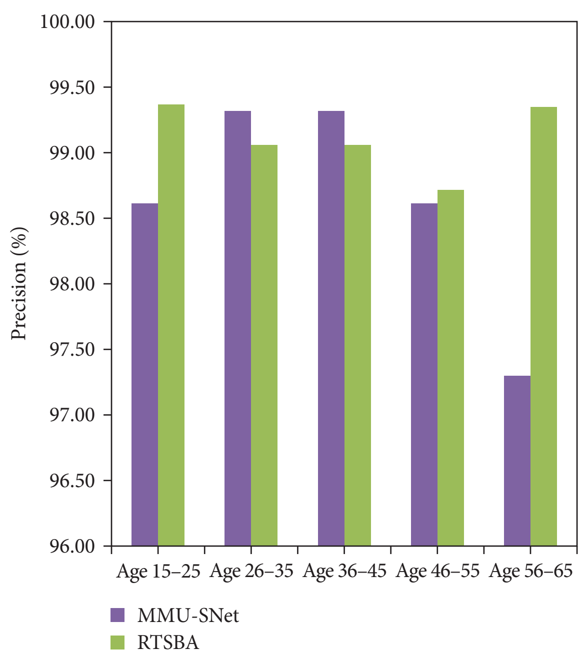


Figure 18: Evaluation of the classifier's performance with different methods through precision. MMU-SNet, modified multi-scale segmentation network.

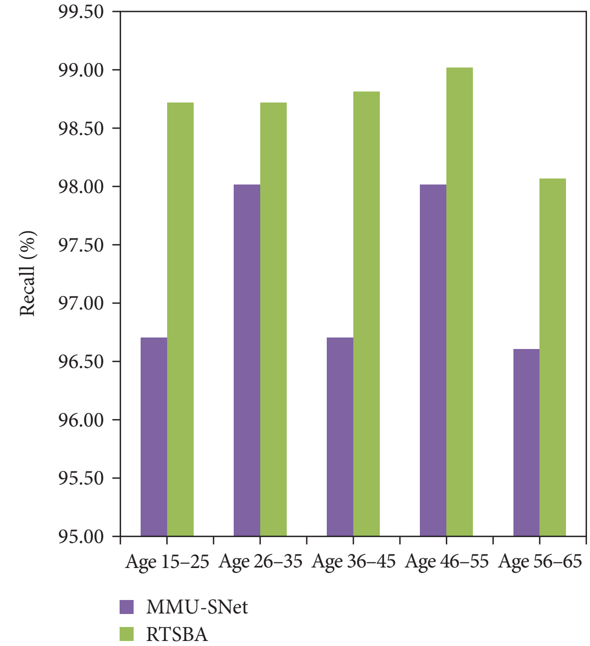


Figure 19: Evaluation of the classifier's performance with different methods through recall. MMU-SNet, modified multi-scale segmentation network.

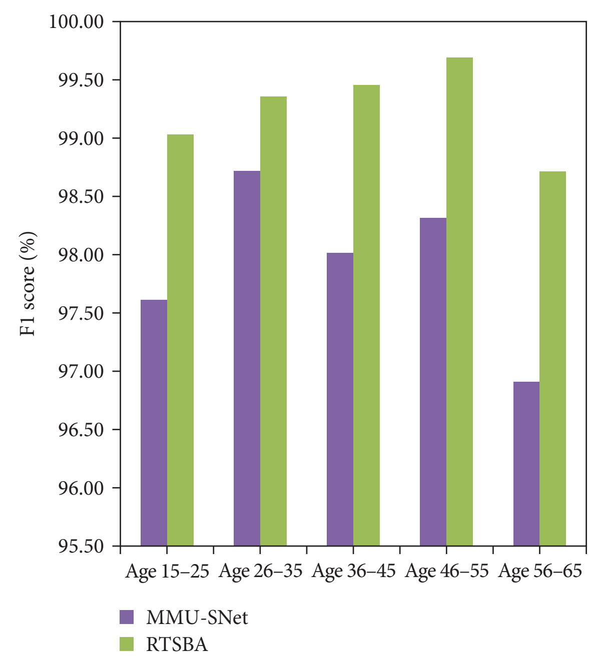


Figure 20: Evaluation of the classifier's performance with the different methods through the F1 score. MMU-SNet, modified multi-scale segmentation network.

complicated connections, and hence it is challenging to figure out the cause behind this behavior. This unpredictability results in performance issues and transient errors being propagated across the time steps, which leads to network instability and causes crashes in the network. RNNs use data from previous time steps; they can fail to use multi-threaded processors, which never support parallelism. Then, the computation time for such networks is higher. CNN is prone to overfitting on the training set. This can lead to poor generalization and inaccurate results on testing data [27]. Compared to a more straightforward signature recognition system, the two-channel and two-stream transformer

Table 8: Various OCR for text conversion.

S. no	OCR types	Recognized accuracy (%)
1.	Google Docs	95
2.	i2 OCR	45
3.	Easy Screen OCR	95
4.	Unicode tamil OCR	35
5.	SUBASA tamil OCR	40
6.	Proposed RTSBA	99

RTSBA, ResNet two-stage bottleneck architecture; OCR, optical character recognition.

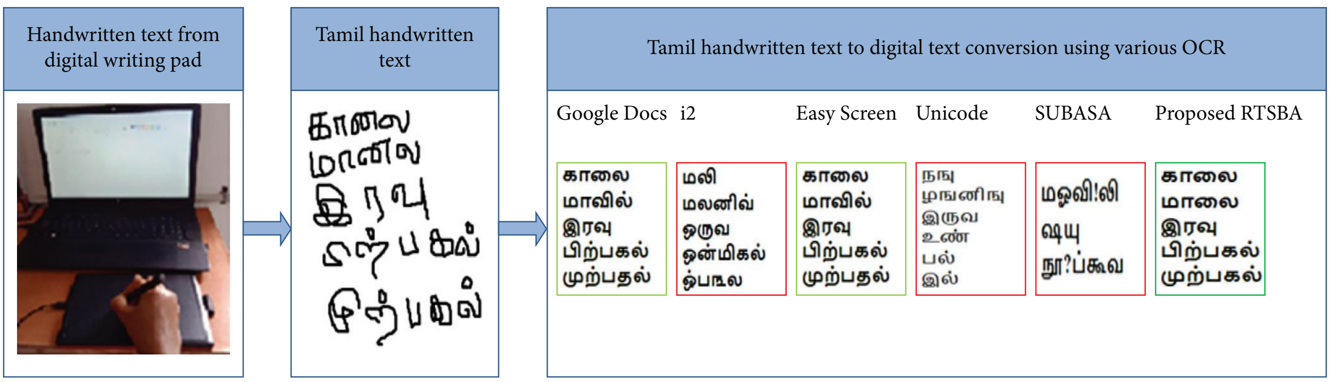


Figure 21: Tamil handwritten text conversion using various OCR.

(2C2S)-based framework [28] requires several additional layers for processing. This increases the amount of processing power for accurate classification. Large amounts of data must be processed, and create an accurate identification system. Data must be kept for a long time when dealing with signatures to ensure accuracy. This leads to high maintenance costs and data storage.

7. Conclusion and Future Work

The RTSBA is proposed for digital writing pad-based handwritten alphabet detection, classification, and word recognition. Handwritten character changes in Tamil words due to pen pressure, pen position, variations in writing styles, spaces between characters of different sizes, unnecessary curves in characters, simple curves, non-simple curves, straight lines, open curves, and closed curves are detected through the proposed RTSBA method. The RTSBA method is compared with different neural network architectures such as LSTM, Inception-v3, RNN, CNN, 2C2S, and MMU-SNet compared to traditional algorithms; the proposed RTSBA methods have a text prediction accuracy of about 98.7%. This model can be applied to other Indian languages, such as Malayalam and Telugu, for text recognition from a digital writing pad.

Data Availability

Data will be made available after a reasonable request from the author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. A. R. Raj, S. Abirami, and S. M. Shyni, "Tamil handwritten character recognition system using statistical algorithmic approaches," *Computer Speech & Language*, vol. 78, Article ID 101448, 2023.

[2] V. Jayanthi and S. Thenmalar, "A review on recognizing offline Tamil manuscript character," *AIP Conference Proceedings*, vol. 2591, no. 1, Article ID 020039, 2023.

[3] A. Fateh, M. Rezvani, A. Tajary, and M. Fateh, "Persian printed text line detection based on font size," *Multimedia Tools and Applications*, vol. 82, pp. 2393–2418, 2023.

[4] K. Shanmugam and B. Vanathi, "Newton algorithm based DELM for enhancing offline tamil handwritten character recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 36, no. 5, Article ID 2250020, 2022.

[5] A. Fateh, M. Fateh, and V. Abolghasemi, "Multilingual handwritten numeral recognition using a robust deep network joint with transfer learning," *Information Sciences*, vol. 581, pp. 479–494, 2021.

[6] Z. Xie, Z. Sun, L. Jin, H. Ni, and T. Lyons, "Learning spatial-semantic context with fully convolutional recurrent network for online handwritten chinese text recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 8, pp. 1903–1917, 2017.

[7] Y. Wang, Y. Yang, W. Ding, and S. Li, "A residual-attention offline handwritten Chinese text recognition based on fully convolutional neural networks," *IEEE Access*, vol. 9, pp. 132301–132310, 2021.

[8] S. Tabassum, N. Abedin, M. M. Rahman et al., "An online cursive handwritten medical words recognition system for busy doctors in developing countries for ensuring efficient healthcare service delivery," *Scientific Reports*, vol. 12, Article ID 3601, 2022.

[9] S. K. Singh and A. Chaturvedi, "Leveraging deep feature learning for wearable sensors based handwritten character recognition," *Biomedical Signal Processing and Control*, vol. 80, Part 1, Article ID 104198, 2023.

[10] A. Srinivas, T.-Y. Lin, N. Parmar, J. Shlens, P. Abbeel, and A. Vaswani, "Bottleneck transformers for visual recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16519–16529, Nashville, TN, USA, 2021.

[11] Sakshi and V. Kukreja, "A retrospective study on handwritten mathematical symbols and expressions: classification and recognition," *Engineering Applications of Artificial Intelligence*, vol. 103, Article ID 104292, 2021.

[12] R. Ghosh, "A recurrent neural network based deep learning model for text and non-text stroke classification in online handwritten Devanagari document," *Multimedia Tools and Applications*, vol. 81, pp. 24245–24263, 2022.

[13] N. Sasipriyaa, P. Natesan, R. Anand, P. Arvindkumar, R. S. Arwin Prakadis, and K. Aswin Surya, "Recognizing handwritten offline Tamil character by using cGAN & CNN," in *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, pp. 509–515, IEEE, Erode, India, 2022.

[14] V. Carbune, P. Gonnet, T. Deselaers et al., "Fast multi-language LSTM-based online handwriting recognition," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 23, pp. 89–102, 2020.

[15] J. Chen, W. Chen, A. Zeb, and D. Zhang, "Segmentation of medical images using an attention embedded lightweight network," *Engineering Applications of Artificial Intelligence*, vol. 116, Article ID 105416, 2022.

[16] A. Ashiquzzaman, A. K. Tushar, A. Rahman, and F. Mohsin, "An efficient recognition method for handwritten Arabic numerals using CNN with data augmentation and dropout," in *Data Management, Analytics, and Innovation*, V. Balas, N. Sharma, and A. Chakrabarti, Eds., pp. 299–309, Springer, 2019.

[17] C. Neche, A. Belaid, and A. Kacem-Echi, "Arabic handwritten documents segmentation into text-lines and words using deep learning," in *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, pp. 19–24, IEEE, Sydney, NSW, Australia, 2019.

[18] R. Pramanik and S. Bag, "Segmentation-based recognition system for handwritten Bangla and Devanagari words using conventional classification and transfer learning," *IET Image Processing*, vol. 14, no. 5, pp. 959–972, 2020.

[19] S. Kundu, S. Paul, S. K. Bera, A. Abraham, and R. Sarkar, "Text-line extraction from handwritten document images using GAN," *Expert Systems with Applications*, vol. 140, Article ID 112916, 2020.

[20] B. Soujanya, S. Chittineni, T. Sitamahalakshmi, and G. Srinivas, "A CNN based approach for handwritten character identification of Telugu Guninthalu using various optimizers," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, 2022.

[21] R. Babitha Lincy and R. Gayathri, "Optimally configured convolutional neural network for Tamil handwritten character recognition by improved lion optimization model," *Multimedia Tools and Applications*, vol. 80, pp. 5917–5943, 2021.

[22] A. Vijay, M. Deepak, P. Kavin, K. Ramar, and B. Shajith, "Transfer learning based offline handwritten recognition system using Tamil characters," in *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, pp. 214–220, IEEE, Erode, India, 2022.

[23] V. Jayanthi and S. Thenmalar, "Recognition of handwritten words from digital writing pad using MMU-SNet," *Intelligent Automation & Soft Computing*, vol. 36, no. 3, pp. 3551–3564, 2023.

[24] S. D. Cui, Y. L. Su, R. Q. dao er ji, and Y. T. Ji, "An end-to-end network for irregular printed Mongolian recognition," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 25, pp. 41–50, 2022.

[25] R. Gayathri and R. Babitha Lincy, "Transfer learning based on handwritten character recognition of Tamil script using the inception-V3 model," *Journal of Intelligent & Fuzzy Systems*, vol. 42, no. 6, pp. 6091–6102, 2022.

[26] R. Ghosh, C. Vamshi, and P. Kumar, "RNN based online handwritten word recognition in Devanagari and Bengali scripts using horizontal zoning," *Pattern Recognition*, vol. 92, pp. 203–218, 2019.

[27] S. Gayathri Devi, S. Vairavasundaram, Y. Teekaraman, R. Kuppusamy, and A. Radhakrishnan, "A deep learning approach for recognizing the cursive Tamil characters in palm leaf manuscripts," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3432330, 15 pages, 2022.

[28] J.-X. Ren, Y.-J. Xiong, H. Zhan, and B. Huang, "2C2S: a two-channel and two-stream transformer based framework for offline signature verification," *Engineering Applications of Artificial Intelligence*, vol. 118, Article ID 105639, 2023.