

An Online Cursive Handwritten Medical Words Recognition System for Busy Doctors in Developing Countries for Ensuring Efficient Healthcare Service Delivery

Shaira Tabassum (✉ shaira.sheba@gmail.com)

Kyushu University

Md Mahmudur Rahman

Global Communication Center, Grameen Communications

Nuren Abedin

Kyushu University

Md Moshir Rahman

Hiroshima University

Mostafa Taufiq Ahmed

Sylhet MAG Osmani Medical College

Ashir Ahmed

Kyushu University

Research Article

Keywords: Smartpen, eHealth, Online Handwriting Recognition, Portable Health Clinic (PHC), Machine Learning, Data Augmentation, Bidirectional LSTM, Medical Term Corpus

Posted Date: October 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-992698/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on March 4th, 2022. See the published version at <https://doi.org/10.1038/s41598-022-07571-z>.

An Online Cursive Handwritten Medical Words Recognition System for Busy Doctors in Developing Countries for Ensuring Efficient Healthcare Service Delivery

Shaira Tabassum^{1*}, Md Mahmudur Rahman², Nuren Abedin¹, Md Moshir Rahman³, Mostafa Taufiq Ahmed⁴ and Ashir Ahmed^{1,2}

^{1*}Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan.

²Global Communication Center, Grameen Communications, Dhaka, Bangladesh.

³Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan.

⁴M A G Osmani Medical College, Sylhet, Bangladesh.

*Corresponding author(s). E-mail(s):
tabassum.shaira.264@s.kyushu-u.ac.jp;

Abstract

Doctors in developing countries are too busy to write digital prescriptions. Ninety-seven percent of Bangladeshi doctors write handwritten prescriptions, the majority of which lack legibility. Prescriptions are harder to read as they contain multiple languages. This paper proposes a machine learning approach to recognize doctors' handwriting to create digital prescriptions. A 'Handwritten Medical Term Corpus' dataset is developed containing 17,431 samples of 480 medical terms. In order to improve the recognition efficiency, this paper introduces a data augmentation technique to widen the variety and increase the sample size. A sequence of line data is extracted from the augmented images of 1,591,100 samples and fed to a Bidirectional LSTM. Data augmentation includes pattern Rotating, Shifting and Stretching (RSS). Eight different combinations are applied to evaluate the

strength of the proposed method. The result shows 93.0% average accuracy (max: 94.5%, min: 92.1%) using Bidirectional LSTM and RSS data augmentation. This accuracy is 19.6% higher than the recognition result with no data expansion. The proposed handwritten recognition technology can be installed in a smartpen for busy doctors which will recognize the writings and digitize them in real-time. It is expected that the smartpen will contribute to reduce medical errors, save medical costs and ensure healthy living in developing countries.

Keywords: Smartpen, eHealth, Online Handwriting Recognition, Portable Health Clinic (PHC), Machine Learning, Data Augmentation, Bidirectional LSTM, Medical Term Corpus

1 Introduction

A global study reports that Bangladeshi physicians spend less than a minute to each primary consultation whereas physicians of Sweden spend 22.5 minutes [1]. This happens due to the massive population and limited number of physicians available for them in developing countries. The ideal doctor to population ratio recommended by World Health Organization (WHO) is 1:1000 while on the contrary, the ratio in Bangladesh is only 0.304:1000 [2]. The consultation time of doctors' includes listening to patients' problems, inspecting test reports, writing a prescription, explaining the patients' condition and giving advises. As they serve a large number of patients in a very short time, they are left with less time for writing a prescription. Thus, the handwriting on the prescription becomes cursive and indecipherable for the patients and pharmacists. As a result, the pharmacists misread the prescriptions due to the similar appearances or sounds of thousands of medication names [3] and end up providing the patients with wrong medicines.

This research conducted an online survey to understand the current state of handwritten prescription usage in the medical practice of Bangladesh. The survey reports that 97.1% of Bangladeshi doctors still generate handwritten prescriptions. According to Bhuiyan et al. [4], incompetency of understanding doctors' handwritten prescriptions is an obstacle for getting quality health services. Difficulty of reading these prescriptions often causes adverse medical consequences such as selecting wrong medicine, improper number of dosage, and even death. National Academies of Science Institute reports that 7,000 deaths occur in US due to the sloppy handwriting of doctors' [5]. Doctors can be trained to write legible prescriptions, but they get a very short period of time to serve each patient.

This paper proposes a machine learning approach to assist the doctors by recognizing cursive handwriting of doctors' and convert them in readable digital prescriptions. The handwriting recognition system consists of several steps: handwritten data collection, data preprocessing for simplification, increasing

data samples using data augmentation, and build a machine learning model for predicting doctors' handwriting.

Bangladeshi prescriptions are a mixture of Bangla and English words with Latin abbreviations of medical terms [4]. A sample of Bangladeshi prescription is given in Fig. 1. Due to the unavailability of enough Bangladeshi prescriptions online, this research started its journey by creating a 'Handwritten Medical Term Corpus'. Initially, a medical words corpus is created using the 8,324 Bangladeshi prescriptions of the Portable Health Clinic (PHC). PHC is a remote healthcare system which provides affordable and sustainable health services to the base of the pyramid population of developing countries [6][7] to ensure universal healthcare coverage targeting rurality, poverty and disability [8]. The PHC data has been used for predicting health status of the existing patients [9][10] in terms of healthcare cost reduction [11], understanding consumer behavior [12]. However, none of these works considered the doctors' role in the PHC system.

The ultimate performance of a remote healthcare system depends on the decision made by the doctor. As mentioned earlier, 97% of the Bangladesh doctors still write handwriting prescriptions. The trend is changing to digital prescriptions. The analog prescriptions are not archived in digital forms and are difficult to search previous medical history. In order to create real time digital prescriptions, a real time handwriting recognition is necessary. At the first step, a 'Medical Term Corpus' is needed. We have created a corpus considering the most used words appeared in the PHC prescriptions. The corpus contains 480 medical-related words (English: 320 and Bangla:120). Afterwards, the handwriting of 39 healthcare professionals of these 480 words are collected for the recognition purpose. Thus, the 'Handwritten Medical Term Corpus' has 17,431 handwritten instances.

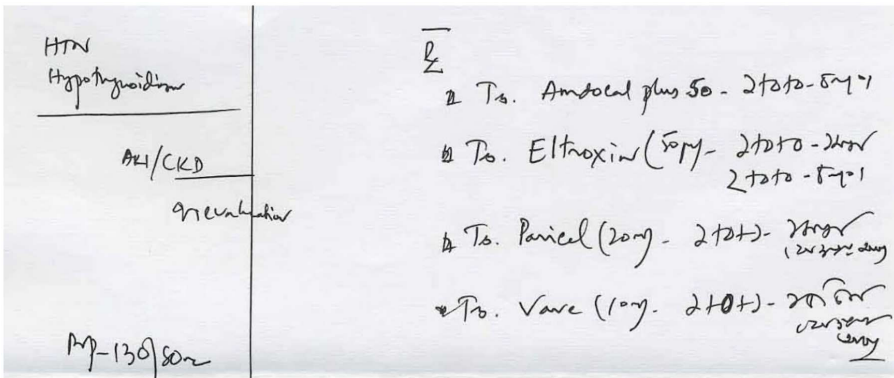


Fig. 1 A sample image of Bangladeshi handwritten prescription

Recognizing different types of handwritings requires a large dataset collected from various sources which is both costly and time consuming. One way to deal with this problem is data augmentation [13]. This paper proposes a

new data augmentation technique - Rotate, Shift and Stretch (RSS) to generate multitudes of handwriting variations. RSS method takes each stroke of a handwritten word and creates new data by updating the coordinates. After applying RSS, the extended dataset has 1,591,100 samples. For predicting the handwritten words, Bidirectional LSTM model is used due to the recent popularity of Recurrent Neural Network (RNN) in the area of handwriting recognition. According to Zhang et al. [14], sequence data can contain rich details of handwriting than static image-like representations.

The proposed handwritten technology can be used in a smart pen, specifically designed for the doctors. The smart pen will digitize the handwriting of doctors into readable texts. The database will store each doctors' writings individually to learn the unique pattern of writing of that particular doctor. This will lead the tool recognize the handwriting of each doctor more efficiently. The smart pen will benefit the doctors by saving time and reducing typographical errors of digital prescriptions.

The rest of the article starts with a review on related researches in Section II, introduces a handwritten dataset of Bangladeshi doctors in Section III, demonstrates the steps of recognition methodology in Section IV, Section V reports the results and findings including the idea of a smartpen, and finally the conclusion is given in Section VI.

2 Related Work

Over the last few decades, multitudes of deep learning approaches have been proposed for efficient handwriting recognition using several handwritten datasets of different languages. This section discusses the similar research works in the following four sectors:

2.1 Doctors' Handwriting Dataset

Few online datasets are available to design a doctors' handwriting recognition system. Dibyajyoti et al. have introduced HP_DocPres dataset with 11,340 samples of handwritten and printed words collected from various medical prescriptions [15]. This dataset is prepared to differentiate between handwritten and printed texts. However, the words are not labeled so it can't be used to recognize the written words by doctors. Another doctors' handwriting dataset is introduced by Farjado et al. [16]. This dataset contains 1,800 images of 12 medicine names collected from 50 doctors from clinics and hospitals of Metro Manila, Quezon City and Taytay, Rizal in Philippines. However, this dataset is not suitable for recognizing doctors' handwriting in Bangladeshi prescriptions due to the limited number of medical terms it contains and the region of data collected being different from our study region.

Although doctors handwriting dataset is scarce, there are multitudes of available handwriting datasets both for English and Bangla languages. IAM Dataset by University of Bern [17] is one of the most popular datasets with the largest handwriting collection in English. This dataset contains 13,353 images

of handwritten lines of text created by 657 writers. A similar dataset in Bangla is the Bangla handwriting recognition dataset by Bappaditya et al. [18] that has obtained 79,000 handwritten Bangla word samples written by 77 different writers. BanglaLekha-Isolated [19] and ISI [20] dataset comes with a vast number of handwriting samples of individual Bangla characters with numerals. Another popular is dataset CMATERdb1 [21] that has 100 handwritten Bangla pages and 50 handwritten English and Bangla combined pages with ground-truth annotations. However, these datasets do not contain any medical terms, hence might perform poorly in recognizing doctors' handwriting.

2.2 Offline Handwritten Character Recognition: using image data as input

Automatic conversion of handwritten texts into images for recognition using Convolutional Neural Network (CNN) is called Offline Character Recognition [22]. Shahariar et al. have proposed a lightweight CNN model for Bangla handwriting recognition. The model has 13 convolutional layers with 2 sub-layers. The sub-layers are joined together to pass through a max-pooling layer with one 0.25 weighted dropout layer. This model has attained 98%, 96.8% and 96.4% accuracy in BanglaLekha, CMATERdb and ISI dataset [23]. A modified LeNet-5 CNN model by Yuan et al. has obtained an accuracy of 93.7% for uppercase and 90.2% for lowercase for the recognition of English language characters [24]. Yang et al. have presented a path-signature feature method using deep CNN for identifying Chinese character writers. The method was 99.52% accurate with DropStroke data augmentation [25].

2.3 Online Handwritten Character Recognition: using time-series data as input

The online character recognition considers a sequence of times which is captured by the movements of a specialized pen. The recognition rate of the online system is more efficient and higher than the offline system [22]. RNN has recently been widely used in the area of handwriting recognition for showing better recognition performance. The RNNs work with sequence data of coordinates which contain vast information than static images [14]. Bappaditya et al. have used bidirectional LSTM using 65,620 handwritten Bangla words dataset and has obtained 79% accuracy [18]. Zhang et al. have proposed a conditional RNN-based generative model combining LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Units). The model is built for recognizing Chinese handwritten characters and has achieved 98.15% recognition accuracy [14]. Farjado et al. have used Convolutional RNN (CRNN) for recognizing doctors' cursive handwriting in English language. The model contains 13 convolutional layers followed by 3 bidirectional LSTM layers and has attained 72% accuracy [16]. However, Achkar et al. has reported obtaining 95% accuracy using the similar CRNN model with a different dataset for recognizing medical handwritten prescriptions [26].

2.4 Handwriting Recognition with Data Augmentation

In our previous work, SRP (Stroke Rotation and Parallel-shift) data augmentation technique was applied for expanding the doctors' cursive handwritten dataset. However, the minimum accuracy of that system was only 68.0% [27]. For recognizing Bangla handwriting characters, Shahariar et al. have applied three data augmentation methods on 10% of the dataset: shifted height and width, rotated images by 10 degrees, and zoomed in the images [23]. Another data augmentation method named 'DropStroke' was used for Chinese character recognition. Chinese characters are very complex as they have many strokes. Thus, the DropStroke method randomly excludes several strokes and generates new handwritten characters using the combination of the remaining strokes [14][25]. Hayashi et al. have used a data augmentation technique using probability distribution for handwriting recognition. This method calculates probability distribution from the features related to the structure of the character. Then, it generates strokes based on the distribution and forms multitudes of new characters [13].

3 Handwritten Medical Term Corpus

This section describes the data collection and preparation process for creating a handwritten medical term corpus.

3.1 Creating medical corpus from digital prescriptions

Medical terms were collected from a remote healthcare prescription database, called PHC (Portable Health Clinic). PHC system maintains an electronic journal of patients' health records. There are major five categories of data in the journal: i) registration data, ii) survey data, iii) clinical data, iv) conversation data, and v) prescription data. The foremost section of 'Handwritten Medical Term Corpus' is collected from the digital prescriptions of PHC. Fig. 2 shows a sample from PHC prescription data.

SL	CC	Drugs				Advices	Test
1	Low back pain,	Tab - Naprosyn - 500 mg - 1+0+1	After meal		5 Days	Previous rules will continue to medication. Do not work long being bent , Use high commode , Take plenty of water , Avoid lifting weights, Please take plenty of water , Do not work long being bent , Avoid lifting weights, Avoid foods that are high in fat and contain higher amount of oil and spices, Please consult with Eye specialist.	S.creatinine ,
		Tab - Beklo - 5 mg - 1+0+1	After meal		5 Days		
2	headache, low back pain, acidity, constipation.	Tab - Neocceptin-R - 150 mg - 1+0+1	Before meal		15 Days		
		Tab - Ace - 500 mg - 1+1+1	After meal		3 Days		
		Tab - Calbo D - - 0+0+1	After meal		Cont.		
3	DM on insulin	Syp - Avolac - - 1 tea spoon	After meal	3 times -- if constipation.		the Diabetic food-chart, Please check blood sugar fasting and before dinner. then adjust the dose of insulin.	

Fig. 2 Sample of digital prescription of PHC

A total of 8,324 digital prescriptions were found in the PHC database. Each prescription contained several columns such as symptoms, medicine names, advices. Initially, a corpus of medical terms is created using the most frequently

appeared words in these columns. The corpus has selected 360 English and 120 Bangla words. These words are sorted according to the frequency of their appearance in the prescriptions, as shown in Fig. 3.

1	426,Ostocal D
2	381,Maxpro
3	376,Ranitid
4	367,Napa
5	322,Ferocit
6	285,Neuro B
7	278,Comet
8	261,Omidon
9	261,Multivit plus
10	261,Colho D

Fig. 3 A segment from the Medical Term Corpus

3.2 Obtaining handwritten data via android application

A simple android app has been developed to obtain real handwriting data from doctors and medical professionals. The app displays medical words on the device screen one by one from the ‘Medical Term Corpus’. Then, the data providers write the corresponding words shown on the screen. Samsung Galaxy Tab S3 was used to capture data as it contains a stylus pen. The data providers write down on the screen using the stylus pen and the application stores the handwritten words in the database. Besides the writings, the application also receives detailed information such as pen movements (xy coordinates) and the status of the pen. Status indicates the state of the pen whether it is up or down. Collected information is stored in the database along with the writings as sequential data. The ‘Handwritten Medical Term Corpus’ also contains the original corpus data to use it as truth value while training the machine learning model. The complete data collection process is shown in Fig. 4.

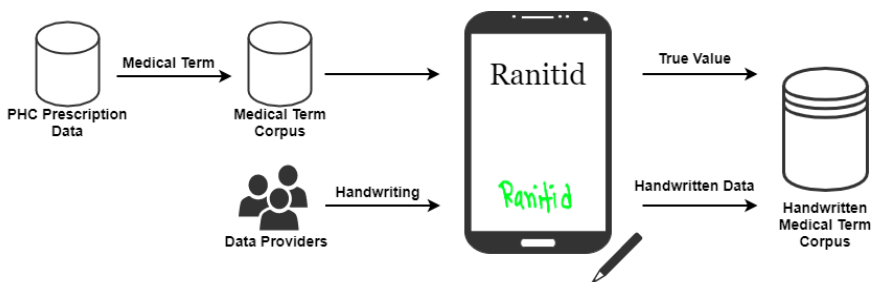


Fig. 4 Handwritten Medical Term Corpus data collection

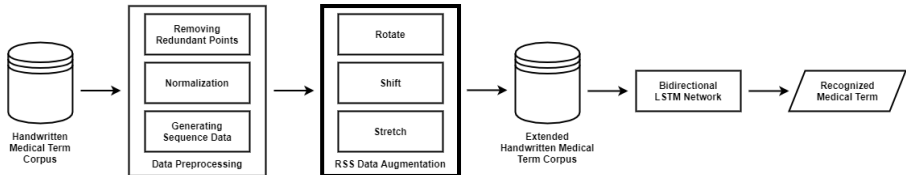


Fig. 5 Overview of the handwritten medical terms recognition system

3.3 Dataset Profile

There are 480 medical words (360 English and 120 Bangla) in the ‘Handwritten Medical Term Corpus’. These words are chosen based on the number of appearances in 8,324 Bangladeshi prescriptions. The hand-writings are collected from 39 medical professionals and doctors of Bangladesh. Due to receiving incomplete data from 12 data providers, 1,289 samples are missing in the dataset. Hence, the dataset has 17,431 handwritten instances of 480 medical-related words. All the data were collected by maintaining authenticity, security and privacy of the data providers and the experiments were performed in accordance with relevant guidelines and regulations.

4 Methodology of Handwriting Recognition

After the data collection phase, the research is administered in three steps, as shown in Fig. 5. First, the collected dataset is analyzed and preprocessed. Then, the proposed RSS (Rotation, Shift and Stretch) data augmentation technique is applied on the preprocessed dataset to expand the number of instances. In this step, a sequence line data is generated from the extended dataset. Finally, bidirectional LSTM model uses the sequence data as input and predicts handwritten medical terms.

4.1 Data Preprocessing

Image-like representations provide very general information about the data. Whereas raw data is rich with information such as spatial and temporal details. These spatio-temporal information can be constituted in a sequence of variable length [14], given in equation (1). Here, x_i and y_i denotes the xy -coordinates of pen movements and s_i states the stroke number of any point n .

$$[[x_1, y_1, s_1], [x_2, y_2, s_2], \dots, [x_n, y_n, s_n]] \quad (1)$$

As shown in Fig. 5, the preprocessing is done in three steps. The images are simplified through removing nearby repetitive points and normalization. Then, a six-dimensional vector is extracted sequentially for each stroke to generate machine learning model input data. The three segments are described below:

4.1.1 Remove Redundant Points

Different styles of handwriting can be found even in the group of people with the same language. Different people follow different way of writing such as small, regular, flat, cursive. Each writer creates distinct sampling points even if they are writing the same character. Thus, a general format for each character can be created by removing nearby redundant points for efficiently estimating the handwritten words. To remove all the redundant points from any handwritten word or character, consider a particular point (x_i, y_i, s_i) where $s_i - 1 = s_i = s_i + 1$. There are two conditions to determine if point i should be removed:

1. Distance between points: If the distance between two points i and $i - 1$ is very small, then point i is removed. In the given equation (2), the threshold $T_{dist} = 0.005 * \max(H, W)$, where H indicates the vertical and W indicates the horizontal widths of the handwriting text input place. Besides, two connecting points lying on a straight line is also be removed.

$$\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} < T_{dist} \quad (2)$$

2. Cosine similarity: Cosine similarity determines the similarity of an inner product space between two non-zero vectors. Point i is removed if similarity between two points i and $i - 1$ is greater than the threshold cosine angle value. In the given equation (3), the threshold T_{cos} is set to 0.99.

$$\frac{\Delta x_{i-1} \Delta x_i + \Delta y_{i-1} \Delta y_i}{(\Delta x_{i-1}^2 + \Delta y_{i-1}^2)^{0.5} (\Delta x_i^2 + \Delta y_i^2)^{0.5}} > T_{cos} \quad (3)$$

Point i is considered as a redundant point if one of the given conditions is satisfied. After removing the redundant points, the shape of the handwritten character is well-preserved and each point of the recreated form contains more information [14]. An example of removing redundant points on a sample handwritten medical term is shown in Fig. 6.

4.1.2 Normalization

After removing the redundant points, the data is normalized for simplification. For x and y coordinates, the maximum x_{max} , y_{max} and the minimum x_{min} , y_{min} is calculated from each data point. Then, x as X and y as Y coordinates is normalized to X_{nor} and Y_{nor} using equation (4). Thus, the (x, y) coordinates data is scaled between the value of 0 and 1.

$$X_{nor} = \frac{X - x_{min}}{x_{max} - x_{min}} \quad (4)$$

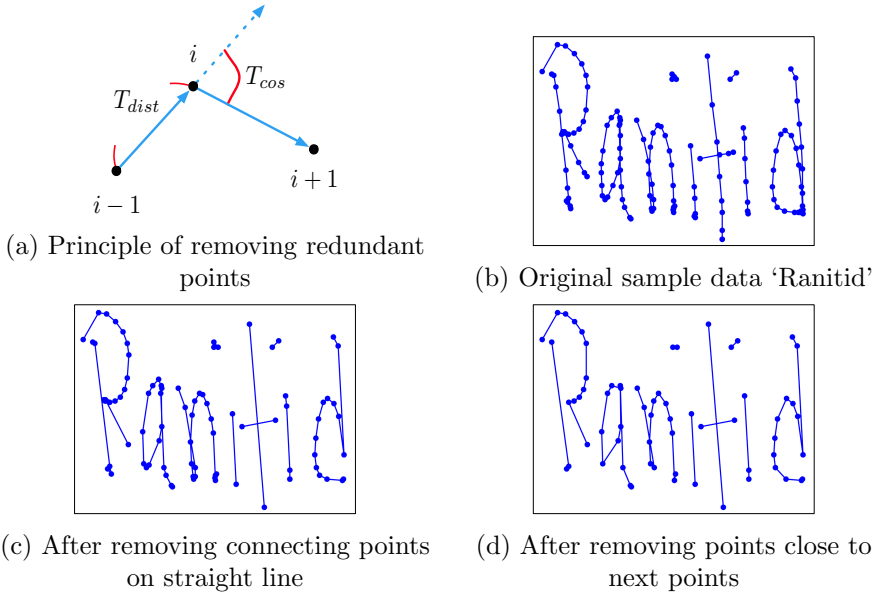


Fig. 6 Remove redundant points from handwritten data

4.1.3 Generate Sequence Data

After the data preprocessing steps, this research has connected the normalized points to form straight lines. Then, a six-dimensional vector is generated from the straight line as L_i with two connecting points i and $i + 1$, as shown in equation (5).

$$L_i = [x_i, y_i, \Delta x_i, \Delta y_i, I(s_i = s_{i+1}), I(s_i \neq s_{i+1})] \quad (5)$$

In the given equation, x_i and y_i are the xy-coordinates which states the starting position of a line. The direction of pen movements in x and y axis is denoted by Δx_i and Δy_i . The last two expressions determine the status of the pen ($[0, 1]$ indicates pen-up, $[1, 0]$ indicates pen-down). The term $I(s_i = s_{i+1}) = 1$ indicates that the starting and ending points of the straight line lies on the same stroke. The last expression $I(s_i \neq s_{i+1}) = 1$ states that the line has moved to the next stroke. Thus, a new sequence of vectors $[L_1, L_2, \dots, L_{n-1}]$ is formed from the (x, y, s) coordinates. This newly generated sequence is denoted as $[x_1, x_2, \dots, x_k]$ for simplification, where each x_i represents one six-dimensional vector [14].

4.2 Data Augmentation

This research has used data augmentation on the preprocessed data in order to increase the number of instances of ‘Handwritten Medical Term Corpus’. A new data augmentation approach named RSS (Rotation, Shift and Stretch) is proposed in this article. RSS method expands data by rotation, shifting

and stretching the shape of character. This method is specifically designed to expand the variety of handwriting styles.

4.2.1 Rotate (Stroke Rotation)

In stroke rotation, the middle point of a stroke (a, b) is determined using the starting point (x_f, y_f) and ending point (x_l, y_l) coordinates, as in equation (6). Then, all the points lying on the middle point of that stroke are rotated. The principle of the rotation process is shown in Fig. 7(a).

$$(a, b) = \left(\frac{x_f + x_l}{2}, \frac{y_f + y_l}{2} \right) \quad (6)$$

Stroke rotation process uses equation (7) to rotate a point (x, y) in θ angle around the middle point of the stroke (a, b) . The rotated point is (X, Y) .

$$\begin{pmatrix} X - a \\ Y - b \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x - a \\ y - b \end{pmatrix} \quad (7)$$

This equation is applied to all of the points on the stroke to rotate the entire stroke around the midpoint of the stroke. Fig. 7(b) is a sample example where the blue color is the original instance and red color is the instance after rotation. This method is applied to the strokes with random values of angles in order to create new augmented data in different forms.

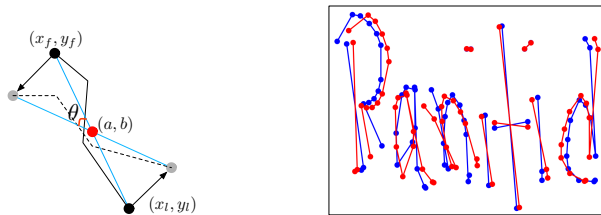


Fig. 7 (a) Principle of Rotation, (b) Sample data after preprocessing (blue) and after Rotation (red)

4.2.2 Shift (Stroke Parallel-shift)

In this second method, all of the points on the stroke is added to a constant number (x, y) in order to shift the strokes in parallel. The principle of parallel-shifting is given in Fig. 8(a). One certain point (x, y) is shifted to a new point (X, Y) following equation (8).

$$(X, Y) = (x + dx, y + dy) \quad (8)$$

The entire stroke is shifted by (dx, dy) after applying this equation to every point on the stroke. Fig. 8(b) is a sample example where the blue color is the

original instance and the red color is the instance after shifting. This method is applied to the strokes with random values of dx and dy in order to create new augmented data in different forms.

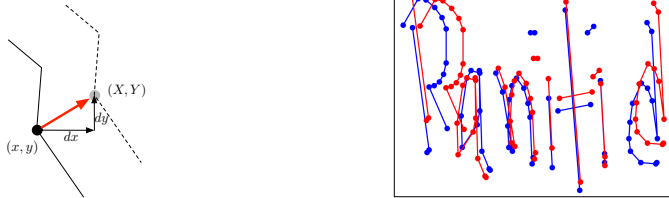


Fig. 8 (a) Principle of Shifting, (b) Sample data after preprocessing (blue) and after Shifting (red)

4.2.3 Stretch (Stroke stretching)

When a writer writes very quickly and roughly, the letters may be stretched vertically or horizontally. Hence, the stroke stretching method is proposed which takes all the strokes of a word and stretches the strokes to change the ratio of the handwritten word. The principle of this method is shown in Fig. 9(a). First, the average value of the y-coordinates are calculated for all the points of a stroke as the reference value Y . If a certain point (x_i, y_i) has larger y-coordinate than Y , then the value of y_i is multiplied by $(1 + r)$, otherwise it's multiplied by $(1 - r)$, shown in equation (9). Here, r is the changing ratio rate and is set to 0.02.

$$y_i = \begin{cases} y_i \times (1 + r), & \text{if } Y < y_i \\ y_i \times (1 - r), & \text{otherwise} \end{cases} \quad (9)$$

However, when the writing is stretched, the points do not overlap and doesn't broke the shape of the character. Fig. 9(b) is a sample example where the blue color is the original instance and the red color is the instance after stretching. According to the findings of this research, this is an effective method for identifying rough handwritings.

4.3 Regenerating Sequence Data

The original dataset contains six-dimensional vectors for all the instances. After data augmentation, the vectors need to be regenerated for the new augmented images by updating the parameters. For the rotated images, the values are updated by adding θ with Δx_i and Δy_i , equation (10). To update the vector for parallel-shifting, dx and dy is added to x_i and y_i , equation (11). Finally, the values of x_i and y_i are multiplied with r for the stretched images, equation (12). However, the values of $I(s_i = s_{i+1})$ and $I(s_i \neq s_{i+1})$ remain unchanged.

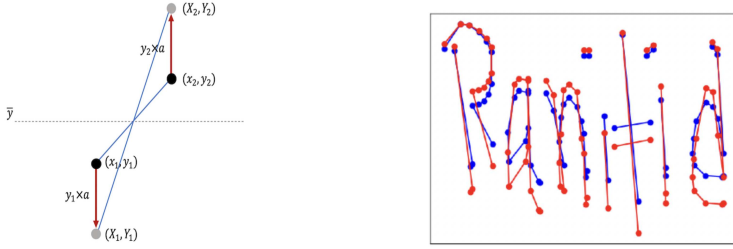


Fig. 9 (a) Principle of Stretching, (b) Sample data after preprocessing (blue) and after Stretching (red)

$$L_i = [x_i, y_i, \Delta x_i + \theta, \Delta y_i + \theta, I(s_i = s_{i+1}), I(s_i \neq s_{i+1})] \quad (10)$$

$$L_i = [x_i + dx, y_i + dy, \Delta x_i, \Delta y_i, I(s_i = s_{i+1}), I(s_i \neq s_{i+1})] \quad (11)$$

$$L_i = [x_i \times r, y_i \times r, \Delta x_i, \Delta y_i, I(s_i = s_{i+1}), I(s_i \neq s_{i+1})] \quad (12)$$

The RSS data augmentation is designed specifically targeting handwriting data as the operations are done by updating the strokes. It can also be used for other datasets if the data are obtained as time series - a sequence of coordinates. If a time series data is converted into images, it can also be used for the expansion of offline characters.

4.4 Machine Learning Model: Bidirectional LSTM

Handwriting contains multiple strokes with several points. Writing style, speed, order, shape of the character varies from person to person which information is difficult to achieve from static images. Hence, this research has dealt with raw sequential data rather than generating image-like representations in order to get rich information about doctors' handwriting.

This research has used Bidirectional LSTM to develop a complete end-to-end recognition system by operating the sequence data extracted from the line data of the augmented handwritten dataset. Bidirectional LSTM uses both past and future inputs for prediction, as shown in Fig. 10, whereas the original LSTM considers only past inputs [28]. In this research, the machine learning model has used both past and future line data to calculate parameters and predict handwritten medical words.

The model architecture for this research is developed using Keras - a neural network library of python. The maximum length of each data is set to 260 by padding zeroes by the end of the instances. There are 300 hidden LSTM layers with corresponding pooling layers. In order to avoid over-fitting, Dropout is used between pooling layer and dense layer [29]. As the model learns the same data many times due to using data augmentation, Early Stopping method is

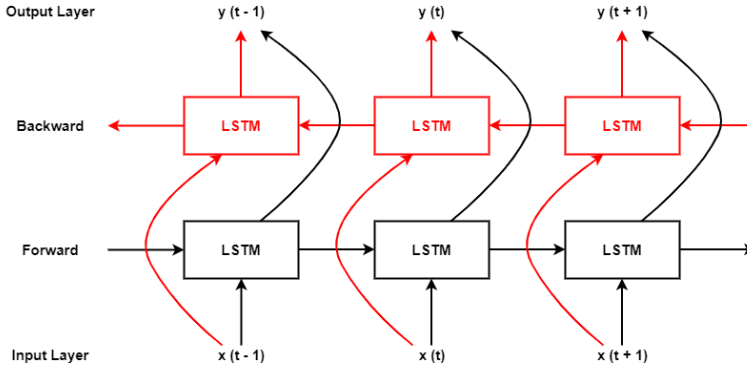


Fig. 10 Concept of Bidirectional LSTM

also used to circumvent overfitting[30]. The Bidirectional LSTM model has the following parameters:

- Activation function: Softmax [31]
- Batch size: 512 [32]
- Learning rate: 0.001 [32]
- Number of epochs: 5 [32]
- Loss function: Categorical cross-entropy [33]
- Optimization function: Adam [34]
- Dropout: 0.3 [29]

5 Results and Discussion

The ‘Handwritten Medical Term Corpus’ contains 17,431 handwritten samples of 480 medical words. The writings are obtained from 39 medical professionals. Among the 39 sets, there are 27 complete sets of instances due to receiving incomplete data from 12 writers. Three sets of 480 words are randomly selected from the complete sets as test data. Thus, the train data has 15,911 and test data has 1,440 handwritten samples. The data augmentation methods are applied only on the train data. Thus, the Bidirectional LSTM model is trained using the extended sequence data and is evaluated based on its performance on the test set.

5.1 Results

This research has performed eight sets of experiments by training the model with eight different combinations of training data. The performance of the model is evaluated based on its accuracy on the same test set. The findings from the model evaluation are as follows, given in Table 1:

1. **No data expansion:** First, the bidirectional LSTM model is trained on the original ‘Handwritten Medical Term Corpus’ with 15,911 instances

Table 1 Data augmentation performance evaluation on Handwritten Medical Term Corpus

SL.	Method	Model Accuracy (%)			Data Size
		Avg	Max	Min	
1.	No data expansion	73.4	90.8	19.0	15,911
2.	Rotate	83.8	95.6	3.33	1,591,100
3.	Shift	91.3	93.9	87.8	1,591,100
4.	Stretch	91.4	95.7	87.5	1,591,100
5.	Rotate + Shift	89.5	95.4	68.0	1,591,100
6.	Rotate + Stretch	90.0	95.9	77.4	1,591,100
7.	Shift + Stretch	88.8	94.2	78.0	1,591,100
8.	Rotate + Shift + Stretch	93.0	94.5	92.1	1,591,100

without applying any data augmentation method. This case has attained the lowest accuracy.

2. **Rotate:** This case has trained by model by applying only Rotation data augmentation method. Rotation is applied 100 times on each instances and the data size becomes 1,591,100. This experiment has achieved high maximum accuracy, but the lowest and unacceptable minimum accuracy which is only 3.33%.
3. **Shift:** Shifting data augmentation is used 100 times on the train data. It has shown overall good performance in all the above three measurements.
4. **Stretch:** Stretching data augmentation is applied similarly on the train data and the expanded data size is 1,591,100. This experiment has achieved similar results to Shifting data augmentation.
5. **Rotate + Shift:** This experiment has combined the Rotate and Shift data augmentation methods. First, each instance is rotated ten (10) times and the data size becomes 159,110. Then, Shift method is performed for another ten (10) times on the expanded instances. Thus, the final data size gets 1,591,100 instances. The maximum accuracy is quite high but the minimum accuracy is comparatively low.
6. **Rotate + Stretch:** The combined Rotation and Stretching methods are applied in this experiment in the similar way. This case has attained the highest maximum accuracy. However, the average and minimum accuracy is low comparing to the other cases.
7. **Shift + Stretch:** Shifting and Stretching data augmentation methods are combined in this case. The data size is 1,591,100 by applying the methods ten (10) times each. This experiment has also accomplished overall considerable accuracy, but better results were observed when these two methods are applied individually on the training data (case **3** and **4**).
8. **Rotate + Shift + Stretch (RSS):** Finally, this is the proposed data augmentation technique of this research. It applies all the three methods on the training data one by one. First, the instances are rotated five (5) times. Then, shifting is performed for another five (5) times on the expanded 79,555 instances. The data size becomes 397,775 where we have applied

the stretch method for four (4) times. Thus, this research introduces the ‘Extended Handwritten Medical Term Corpus’ which contains 1,591,100 handwritten medical term samples. The Bidirectional LSTM model with RSS data augmentation has achieved the highest average and minimum accuracy. It could not reach the highest maximum accuracy but the accuracy never plunged under 92.1%. Thus, this experiment can be considered as the best possible method for recognizing doctors’ cursive handwriting.

5.2 IoT Smartpen for Doctors: An Application Example

The handwriting recognition tool can be installed in a doctor’s smart pen, which is an ongoing work at our research institution. As shown in Fig. 11, the smart pen has five major modules. The same recognition tool can also be installed in tablet PCs. The **pointer** contains a sensor in the nib of the pen to collect words written by a doctor. The **fingertip sensor** will recognize the authorized user of the pen so that unauthorized person can not use the pen. The **memory** will store all the prescriptions so that the doctor can easily find previous health records. It can also store new patterns of words written by the doctor.

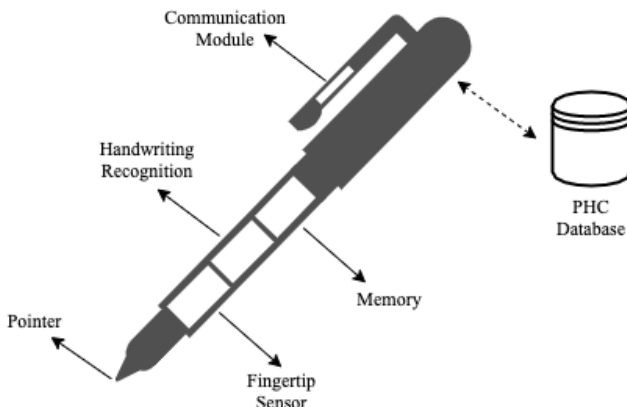


Fig. 11 IoT smart pen design for doctors

In this way, the data set will get larger day by day. The system will also be able to capture the unique handwriting habit of that particular doctor. The **handwriting recognition** tool will recognize the words written by the doctor and convert them into text to store it to the memory. A copy of the prescription will also be stored in **PHC DB** that can be accessed by the patient, authorized pharmacy and family members. The **communication module** will be used as a sound recorder to generate prescriptions from doctors’ speech through voice recognition. However, that module will use a different technology which is not part of this particular research of handwriting recognition.

6 Conclusion

The objective of this research has been to recognize doctors' handwriting and digitize the prescriptions in real time. Towards this goal, this paper contributes in three steps- (a) develop a medical term corpus (b) introduce a unique data augmentation technique and (c) use a machine learning approach for final recognition. It also compares the recognition accuracy in different augmentation stages.

The machine learning approach was designed for recognizing particularly doctors' cursive handwriting and converting them into digital printed texts. A dataset named 'Handwritten Medical Term Corpus' was created from digital prescriptions of PHC that contains 17,431 handwritten texts of 480 Bangla and English medical-related words. A data augmentation method RSS was proposed for enriching the variety of doctors' handwriting. RSS method expanded the data set to 1,591,100 instances which was also introduced in this paper named 'Extended Handwritten Medical Term Corpus'. Bidirectional LSTM model was used to create an online character recognition system for predicting doctors' handwriting. This research performed eight experiments on the handwritten data set and achieved 93.0% average accuracy (max: 94.5%, min: 92.1%) using Bidirectional LSTM and RSS data augmentation. This accuracy was 19.6% higher than the recognition result with no data expansion.

The current accuracy needs to be improved. The proposed recognition methodology can be implemented in a smart pen for doctors. A brief system architecture of the proposed smart pen is introduced. Doctors will use the smart pen for writing and the tool will automatically convert the handwriting texts into digital prescriptions. More samples will be collected and will be automatically stored in the corpus. Doctors personal handwriting habit will also be captured. More samples will improve the recognition accuracy.

Acknowledgement

The authors are grateful to the team in Global Communication Center (GCC) in Grameen Communications, Bangladesh for collecting data samples for creating the 'Medical Term Corpus'. Kaze Shindo and Ryo Takahashi of our Social Tech Lab conducted the data augmentation experiments.

Declarations

- Funding
- Conflict of interest/Competing interests (check journal-specific guidelines for which heading to use)
- Ethics approval
- Consent to participate
- Consent for publication
- Availability of data and materials
- Code availability

- Authors' contributions

References

- [1] Irving, G., Neves, A., Dambha-Miller, H., Oishi, A., Tagashira, H., Verho, A., Holden, J. (2017). International variations in primary care physician consultation time: a systematic review of 67 countries. *BMJ open*, 7(10), e017902.
- [2] Kumar, R., Pal, R. (2018). India achieves WHO recommended doctor population ratio: A call for paradigm shift in public health discourse!. *Journal of family medicine and primary care*, 7(5), 841.
- [3] Jeetu, G., Girish, T. (2010). Prescription drug labeling medication errors: a big deal for pharmacists. *Journal of Young Pharmacists*, 2(1), 107–111.
- [4] Bhuiyan, B., Urmi, I., Chowdhury, M., Rahman, T., Hasan, A., Simkhada, P. (2019). Assessing whether medical language is a barrier to receiving healthcare services in Bangladesh: an exploratory study. *BJGP open*, 3(2).
- [5] Brits, H., Botha, A., Niksch, L., Terblanché, R., Venter, K., Joubert, G. (2017). Illegible handwriting and other prescription errors on prescriptions at National District Hospital, Bloemfontein. *South African Family Practice*, 59, 1-4.
- [6] Ahmed, A., Inoue, S., Kai, E., Nakashima, N., Nohara, Y. (2013). Portable Health Clinic: A pervasive way to serve the unreached community for preventive healthcare. In *International Conference on Distributed, Ambient, and Pervasive Interactions* (pp. 265–274).
- [7] Ahmed, A., Rebeiro-Hargrave, A., Nohara, Y., Maruf, R., Ghosh, P., Nakashima, N., Yasuura, H. (2015). Portable health clinic: A telehealthcare system for unreached communities. In *Smart sensors and systems* (pp. 447–467). Springer.
- [8] Podder, K., Tabassum, S., Khan, L., Salam, K., Maruf, R., Ahmed, A. (2021). Design of a Sign Language Transformer to Enable the Participation of Persons with Disabilities in Remote Healthcare Systems for Ensuring Universal Healthcare Coverage. In *2021 IEEE Technology Engineering Management Conference*.
- [9] Tabassum, S., Sampa, M., Islam, R., Yokota, F., Nakashima, N., Ahmed, A. (2020). A Data Enhancement Approach to Improve Machine Learning Performance for Predicting Health Status Using Remote Healthcare Data. In *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)* (pp. 308–312).

- [10] Sampa, M., Hossain, M., Hoque, M., Islam, R., Yokota, F., Nishikitani, M., Ahmed, A. (2020). Blood Uric Acid Prediction With Machine Learning: Model Development and Performance Comparison. *JMIR medical informatics*, 8(10), e18331.
- [11] Tabassum, S., Sampa, M., Maruf, R., Yokota, F., Nakashima, N., Ahmed, A. (2020). An Analysis on Remote Healthcare Data for Future Health Risk Prediction to Reduce Health Management Cost. In *APAMI 2020: 11th Biennial Conference of the Asia-Pacific Association for Medical Informatics* (pp. 115-119).
- [12] Hossain, M., Okajima, H., Kitaoka, H., Yokota, F., Ahmed, A. (2018). eHealth Consumer Behavior. In *Behavior Engineering and Applications* (pp. 129–147). Springer.
- [13] T. Hayashi, K. Gyohten, H. Ohki, T. Takami (2018). A Study of Data Augmentation for Handwritten Character Recognition using Deep Learning. In *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (pp. 552-557).
- [14] Zhang, X.Y., Yin, F., Zhang, Y.M., Liu, C.L., Bengio, Y. (2017). Drawing and recognizing chinese characters with recurrent neural network. *IEEE transactions on pattern analysis and machine intelligence*, 40(4), 849–862.
- [15] Dhar, D., Garain, A., Singh, P., Sarkar, R. (2020). HP_DocPres: a method for classifying printed and handwritten texts in doctor’s prescription. *Multimedia Tools and Applications*, 1–34.
- [16] Fajardo, L., Sorillo, N., Garlit, J., Tomines, C., Abisado, M., Imperial, J., Rodriguez, R., Fabito, B. (2019). Doctor’s Cursive Handwriting Recognition System Using Deep Learning. In *2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)* (pp. 1–6).
- [17] Bunke, H. (2000). IAM Handwriting Database. Research Group on Computer Vision and Artificial Intelligence, INF, University of Bern.
- [18] Chakraborty, B., Mukherjee, P., Bhattacharya, U. (2016). Bangla Online Handwriting Recognition Using Recurrent Neural Network Architecture. In *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*. Association for Computing Machinery.
- [19] Mithun Biswas, Rafiqul Islam, Gautam Kumar Shom, Md. Shopon, Nabeel Mohammed, Sifat Momen, Anowarul Abedin (2017). BanglaLekha-Isolated: A multi-purpose comprehensive dataset of Handwritten Bangla Isolated characters. *Data in Brief*, 12, 103-107.

- [20] U. Bhattacharya, B. B. Chaudhuri (2009). Handwritten Numeral Databases of Indian Scripts and Multistage Recognition of Mixed Numerals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(3), 444-457.
- [21] Sarkar, R., Das, N., Basu, S., Kundu, M., Nasipuri, M., Basu, D. (2012). CMATERdb1: a database of unconstrained handwritten Bangla and Bangla-English mixed script document image. *International Journal on Document Analysis and Recognition (IJ DAR)*, 15(1), 71-83.
- [22] Priya, A., Mishra, S., Raj, S., Mandal, S., Datta, S. (2016). Online and offline character recognition: A survey. In 2016 International Conference on Communication and Signal Processing (ICCSP) (pp. 0967-0970).
- [23] Rabby, A., Haque, S., Islam, S., Abujar, S., Hossain, S. (2018). Bornonet: Bangla handwritten characters recognition using convolutional neural network. *Procedia computer science*, 143, 528-535.
- [24] Yuan, A., Bai, G., Jiao, L., Liu, Y. (2012). Offline handwritten English character recognition based on convolutional neural network. In 2012 10th IAPR International Workshop on Document Analysis Systems (pp. 125-129).
- [25] Yang, W., Jin, L., Liu, M. (2015). Character-level Chinese writer identification using path signature feature, dropstroke and deep CNN. arXiv preprint arXiv:1505.04922.
- [26] Achkar, R., Ghayad, K., Haidar, R., Saleh, S., Al Hajj, R. (2019). Medical Handwritten Prescription Recognition Using CRNN. In 2019 International Conference on Computer, Information and Telecommunication Systems (CITS) (pp. 1-5).
- [27] Tabassum, S., Takahashi, R., Rahman, M., Imamura, Y., Sixian, L., Rahman, M., Ahmed, A. (2021). Recognition of Doctors' Cursive Handwritten Medical Words by using Bidirectional LSTM and SRP Data Augmentation. In 2021 IEEE Technology Engineering Management Conference-Europe (TEMSCON-EUR) (pp. 1-6).
- [28] Schuster, M., Paliwal, K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11), 2673-2681.
- [29] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15, 1929-1958.
- [30] Prechelt, L. (2012). Early Stopping — But When?. In *Neural Networks: Tricks of the Trade: Second Edition* (pp. 53-67). Springer Berlin

Heidelberg.

- [31] Chigozie Nwankpa, W. Ijomah, A. Gachagan, S. Marshall (2018). Activation Functions: Comparison of trends in Practice and Research for Deep Learning. ArXiv, abs/1811.03378.
- [32] Nagaoka, T. (2020). Hyperparameter Optimization for Deep Learning-based Automatic Melanoma Diagnosis System. *Advanced Biomedical Engineering*, 9, 225–232.
- [33] Zhilu Zhang, Mert R. Sabuncu. (2018). Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels.
- [34] Diederik P. Kingma, Jimmy Ba. (2017). Adam: A Method for Stochastic Optimization.