Kannada Handwritten Answer Recognition using Machine Learning Approach

Sampathkumar S (sampath1447@gmail.com)
Kuvempu University

Research Article

Keywords: pre-processing, segmentation, classification, word recognition

Posted Date: May 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1661060/v1

Kannada Handwritten Answer Recognition using Machine Learning Approach

Ravikumar M[1*], Sampathkumar S[1†] and Shivakumar G[1†]

[1*]Department of Computer Science, Kuvempu University, Shankaraghatta, Shivamogga, 577451, Karnataka, INDIA.

*Corresponding author(s). E-mail(s): ravi2142@yahoo.co.in;
Contributing authors: sampath1447@gmail.com;
g.shivakumarclk@gmail.com;
†These authors contributed equally to this work.

Abstract

Handwriting recognition has been an issue of concern for many researchers and analysts throughout the previous few decades. Different applications need solution to recognize the cursive nature of handwritten text. The stated nature of written styles needs to implement. To build an efficient working OCR the main drawback is to preprocess noisy document, segment the word, character and then recognize the written text. This paper comprises the needs, relevant research towards handwritten recognition and how to process. We line-out the steps and stages used in the recognition of Kannada handwritten words. The main aim of proposed work is to identify Kannada handwritten answer written in answer booklet and to solve recognition problem by using machine learning algorithms. System provides a detailed concept on pre-processing, segmentation, classifier used to develop systematic OCR tool. The achieved accuracy is of 90% for Kannada handwritten words.

Keywords: pre-processing, segmentation, classification, word recognition

1 Introduction

Although, a great success rate has already been achieved in recognition of automatic printed text. In the current era there is no high accuracy gained

for handwritten word recognition. From past duration of time handwritten recognition is an active area in document image analysis, it has vast potential applications. Some of the areas which relates to understanding of text are postal automation, cheque processing, answer script evaluation, automatic form filling in data entry etc... the main drawback behind the problems to construct an robust recognition system is character may be elongated, swooped, slanted, stylized, crunched, connected tiny, gigantic etc... manually transcribing large amounts of handwritten data is arduous process that's bound to be fraught with errors, automated handwritten recognition can drastically take less to convert higher data to text and also provide a architecture for future applications. Recognition of handwritten text is an evolving field of research that surround artificial intelligence, computer vision and pattern recognition. There are two types of handwriting systems online and offline. To progressively implement the applications both types can be used based on user's response. Several methods have been used for the online and offline in the field of handwritten recognition [3]. Now also text recognition is considered as a challenging problem statement. The high variance in handwriting styles across people and poor quality of handwritten text compared with printed text pose relevant hurdles in converting it to readable form and to recognize. It is an acute problem to solve the many industries like health care, insurance and banking sector [4]. To get a desirable result in a word recognition phase, a robust and efficient pre-processing and segmentation is needed. Only proper segmentation will improve the performance of OCR. Without segmentation we cannot directly build an efficient word recognition OCR. Segmentation of Kannada handwritten word into character is also one of the tedious and ongoing challenging task from few decades and as it contains vowels and consonants and apparent mixture of overlapped Ottaksharas, Vyanjanas because of its structured complexity as compared with Latin based languages, Kannada handwritten text is difficult to segment. Moreover, Kannada language has 18511 distinct characters [2], segmentation of handwritten Kannada script into lines, words and character is of great importance and much demanded for some specific applications. Segmentation of handwritten Kannada script poses major challenges due to additional modifier characters.

CHARACTERISTICS OF KANNADA SCRIPT

Kannada is the one of the Dravidian language of India and the official administrative language of the state Karnataka. Karnataka is a linguistically diverse state where Kannada is spoken commonly by the people. It has its own script derived from Brahmi script. Kannada script has base set of 49 characters compromising 15 vowels and 34 consonants. Further there are distinct symbols that modify the base consonants called modifiers [2]. Recognition of Kannada characters is very difficult because of its shape, non-uniqueness, one or more ottaksharas present below the character, drastic variation in handwritten styles. Example of Kannada handwritten document, answer script are

shown in fig 1.

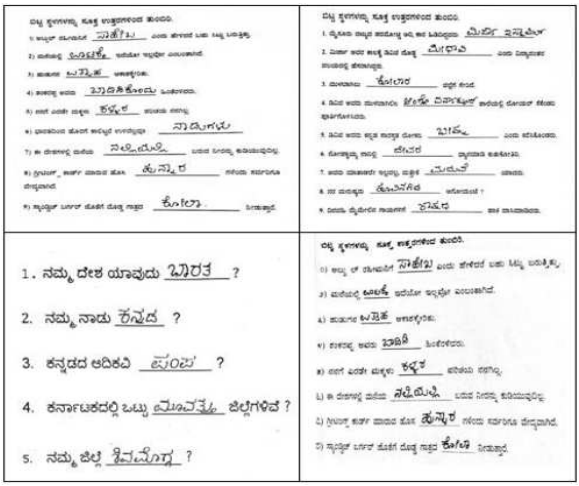


Fig. 1 Samples considered for experimentation.

CHALLENGES IN KANNADA HANDWRITTEN SCRIPT

1. Person to person huge changes and ambiguity of strokes and handwritten styles.
2. Due to inconsistent time to time handwritten font may vary.
3. Due to degradation ink smudges present in the document over time.
4. It is difficult to recognize source document because of poor quality.
5. Recognition of cursive handwriting makes separation and it is tedious to analyze.
6. Printed text always easy to understand and to recognize for the system where all the character sits up straight but in the case of handwritten document text can be invariant, different rotation.

The paper is organized in the following sections: Section 2 contains the review of related work. Section 3 gives brief idea of proposed method used. Section 4 briefs results and discussion and section 5 explains the conclusion of the research work.

2 Related Work

In this review paper [1], Kannada district names recognition is focused using classifier. The database is created by 60 different writers with district names of Karnataka. Euclidean distance and DTW algorithm is used to recognize the word images. Pre-processing stage includes binarization, noise removal, skew detection and correction, thinning. The recognition accuracy provides

considerable results. Zone wise horizontal and vertical profile based method is proposed [2], dataset consists of 750 images of Kannada words. Pre-processing and segmentation is done for the purpose of word recognition neural network is used for training the network. Generated features are used to train the model. Thinning, bounding box generation and feature extraction is done to create the zones. The system works in two stages training and testing. experiments give recognition accuracy of 97.1%.

Locality preserving projections method is proposed [3], system design contains data acquisition, pre-processing, feature extraction classification and recognition. The experiment was conducted on white sheet and images are scanned. For the purpose of proposed algorithm very small network of 5 nodes is used. Algorithm provides very well efficient transmission of data. The accuracy gained is 80%. Convolutional neural network model is proposed [4], proposed system works for two stages training and testing. In the first step input document is converted to colored image to gray scale. In the step 2, denoising is processed to remove noise. In the step 3, pre-processing is done to extract features and to configure CNN model, web application is developed. Efficacy for char74k is 99% and for own created dataset is 96%.

Online recognition of Kannada characters is proposed [5], recognition of words for top, middle and bottom strokes. Data is collected by different age groups of Karnataka. Genius mouse pen was used to collect dataset. These are validated, segmented and features were extracted. Character segmentation is done by individually dividing into top strokes, middle stroke and bottom stroke. For recognition KNN and SVM pattern is used. The maximum yielding rate of accuracy is 92.5% for KNN and 94.35% for SVM.

HMM based character recognition method is proposed [6], the stages involved to develop a model is pre-processing to enhance and render the suitable data. Segmentation to separate individual character feature extraction to store only required feature for recognition. Training is carried out by baumwelch re-estimation procedure. The model is trained for 100 training samples. accuracy of Kannada characters is 76%. Grid based approach method is proposed [7], 28 district names of Karnataka state. Each written from 40 different people is collected for experimentation. The proposed method segregates the input word into four grids. PCA is applied for better representation. For classification distance measure technique is applied and obtained result is 68%.

Template matching method for recognition on correlation analysis is proposed [8]. In the step1 the dataset is collected by test images of many writers and passed for pre-processing stage to process the image. In the step 2 images is segmented into individual characters using simple segmentation method. The segmented images correlated with the already stored images. The maximum related images is showed in the editable format. Hybrid features for recognition of handwritten Kannada characters is proposed [9],3600 samples are collected. K- nearest neighbor classifier is used to classify characters. Proposed techniques uses local and global features. Features are extracted for each individual image. Overall accuracy gained is 87.33%.

In this review paper [10], Kannada character segmentation algorithm is presented. The proposed method is based on the thinning, branch points and mixture models. From touching lines characters are segmented, the proposed method works good for handwritten Kannada words and has shown very well results up to 85.5%. A connected component analysis algorithm is proposed [11], segmentation of handwritten characters and feature extraction is carried out by using feature extraction algorithms. For recognition and classification of characters' support vector machine classifier is used. Experiments shows that an improved in correctness and validity for segmentation of characters.

Projection profiles and morphological operations methods is proposed [12], Kannada handwritten scripts is segmented into lines, words and typescript. In pre-processing phase organizing and disintegration is done. The main goal is to concentrate on singular content line, the experiments is done on absolutely unconstrained Kannada handwritten scripts. Recognition based on zoning is proposed [13], feature extraction method is used for character recognition and support vector machine classifier is used for classification. Training phase and testing phase is to compare characters. In the pre-processing stage cropping, resizing and noise removal is done. Median filter is used to enhance the image. In the segmentation stage ON pixel and OFF pixel is taken as X and X1, X and X2 to find all the characters in a word. At last classifier is applied which increases the speed of the accuracy and which in requires less memory to store training samples. A twelve directional feature extraction technique is proposed [14], data is collected from 25 different writers having different handwritten styles. A feed forward back propagation neural network is used for classification and recognition of Kannada handwritten characters. Morphology and projection technique is to segment the characters. Experimentation is carried out feature extraction method and classifier gives better accuracy rate. Hidden markov model method is proposed [15], implicit segmentation technique is used to avoid explicit segmentation. To segment characters and to reduce number of classes in a characters HMM method is used. Experiment is conducted on 74k char dataset, to train character gradient based feature is extracted. Water reservoir and thinning concept is proposed [16], connected component analysis(CCA) algorithm is to detect the components in a word. Based on touching position, close loop positions and morphological structure of touching region, the character cutting path is generated to segment characters. Proposed concept holds best results. Skew detection, correction and segmentation method is proposed [17], bounding box technique is applied to skew angle detection and correction. It is a way to find the extreme corners of text image. Skew correction is done by rotating the document through an angle with respect to horizontal line. Unconstrained handwritten Kannada script is segmented into lines and words. For line segmentation sobel edge detector is applied which lists the points in an image. Experiment is conducted on 40 document images an average segmentation rate 0f 91% and 70% for lines and words is obtained.

Horizontal projection profile and widowing method is proposed [18], to segment and recognize feedback-based approach is applied. The gap between

segmentation and recognition phase is fixed by attempting a proposed method. KHTD standard dataset is used and got recognition accuracy of 95.02%. Major headings should be typeset in boldface with the first letter of important words capitalized.

3 PROPOSED METHOD

In this section, we discuss the proposed method for line detection, handwritten answer extraction, character segmentation and word recognition. Scanned Kannada handwritten answer scripts are collected for automatic answer script evaluation process which involves image acquisition using scanner.

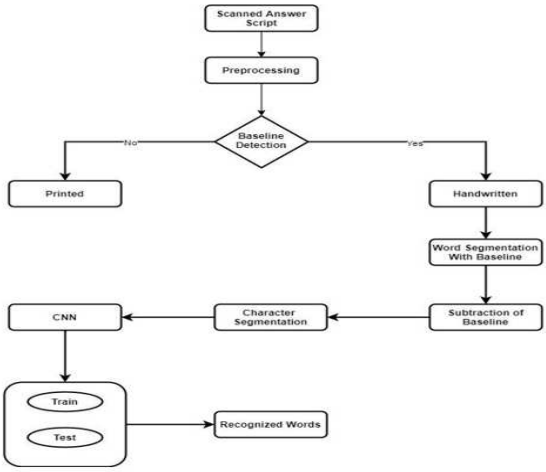


Fig. 2 Block diagram of Proposed Method.

3.1 PRE-PROCESSING

In the step 2, pre-processing is done which contains certain stages to process the document, pre-processing is a stage to eliminate the unabiguties and to remove the uninformative variations in handwriting styles. The main aim of pre-processing an answer script is to improve the image data that suppress unwanted distortions like ink smudges, noise present while using scanner and to enhance some images features important for further processing, depending on the answer script. Some pre-processing we used are:.

3.2 Smoothing

The goal is to smooth the data sets by dividing them into individual segment of the same size. The Gaussian function for smoothing is

$$g(x,y) = \frac{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)f(x+s,y+t)}{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)} \tag{1}$$

where w (s, t) denotes the mask elements, a = (m 1)/2, and b = (n 1)/2. If the sum of the mask elements is more or less than one, the result will be brightened or darkened.

3.3 Dilation

It is a morphological image processing operations. It is a procedure for modifying to geometric structure. Dilate operation is used to remove noise, identify intensity. The operation is performed by using the cv2.dilate() method.

3.4 Erosion

Dilation is very similar to erode, the neighbour pixels with specified window is examined. Dilation function is applied to remove shades where the handwritten document contain noise. The operation is carried out by using cv2.erode().

3.5 Min and Median

Min method is used to thicker the letter of characters. If the text is lightly written on a document. The function is bitmap. transforms. Minimum (2). Median is most common and fruitful image transformation method. It is used to preserve the edges while removing noise.

3.6 Skew and Slant correction

Skew detection and correction in an image processing became a mandatory in the field of DIP and pattern recognition which takes important role in automation of documentation. Slant is the angle in degrees clockwise from vertical at which the characters were drawn. It is process of the characters in a line or document. Skew and slant is corrected and sent to further stage to get better results.

3.7 Augmentation

Text enrichment involves augmenting real text data with information that we processed previously. augmentation provides more semantics original text, hereby enriching its predictive power and the depth of analysis that can perform on data.

3.8 Binarization

A highly overlooked pre-processing step is text normalization. Text normalization is the process of transforming a text into a standard form. For example, the word "good" and "gud" can be transformed to "good" its canonical form. Text normalization is important for noisy text such as social media comments, text messages and blog spots where abbreviations, misspellings are odd out can be removed. In the step 3, straight line detection method is applied to detect baseline, a straight line can be represented as y = mx+b and graphically plotted for each pair of image points (x, y). the straight line is represented using polar parameters r and as r= x cos +y sin. Horizontal line detection is applied to extract exact handwritten answers and word is segmented with baseline from answer script. In the step 4, subtraction of baseline by using line detection method is used which helps to segment word and after removing the baseline, the word is passed to further stage i.e. segmentation.

4 SEGMENTATION

Segmentation of a document image into its basic entities namely text lines and words, is a critical stage towards handwritten document recognition. The difficulties that arise in handwritten documents make the segmentation procedure a challenging task. For the text line segmentation procedure these include the difference in the skew angle between lines on the page or even along the same text line, overlapping words and adjacent text lines touching.

Furthermore, the frequent appearance of accents in many languages (e.g. French, Greek) makes the text line segmentation. For the word segmentation process, difficulties that arise include the appearance of slant in the text line, the existence of punctuation marks along the text line and the non-uniform spacing of words [7]. Perfectly horizontal text lines are trivial to detect using the Vertical Projection Profile (VPP)[12] of an image. The VPP is obtained by projecting the image onto the vertical axis, which effectively is to take the sum of the pixel values in each row.

$$\mathrm{VPP}(f(x,y),w) = \begin{bmatrix} \sum_{i=0}^{w} f(0,i) & \dots & \sum_{i=N-1-w}^{N-1} f(0,i) \\ \vdots & \ddots & \vdots \\ \sum_{i=0}^{w} f(M-1,i) & \dots & \sum_{i=N-1-w}^{N-1} f(M-1,i) \end{bmatrix} \quad (2)$$

The width w is chosen small enough to linearly separate the text lines that they contain, and large enough to render peaks in the VPP. The VPPs are then stored in a 32-bit image with M rows and N columns.

4.1 Word Segmentation

word segmentation is one of the pre-processing step for any handwritten document. The main aim of word segmentation is to extract whole word image in a

handwritten document. It is necessary that word segmentation is very important in a retrieval system. To build a proper character segmentation and to get a good result, line and word segmentation is necessary. There are several segmentation methods available in literature such as projection profile method, connected component analysis etc. The word segmentation is performed by considering the Horizontal Projection Profile (HPP) of the text line image.

$$\mathrm{HPP}(f(x,y)) = \left[\sum_{i=0}^{M-1} f(i,0) \ldots \sum_{i=0}^{M-1} f(i,N-1)\right] \tag{3}$$

A Smoothed HPP is obtained by smoothing the HPP using an averaging filter of width wf, chosen so that the gaps within the words are closed but leaves the gaps between the words open.

4.2 Character Segmentation

Several segmentation techniques are proposed for achieving high recognition accuracy. Recognition accuracy depends on the correct segmentation. In this paper, methods have been proposed to extract answer and to segment an answer scripts into sentences, words and then finally characters for isolated character. Further methods have been used to pre-process the image in such a way that reduces the memory needed by the models to process the characters.

4.3 CONVOLUTIONAL NEURAL NETWORK

One of the supervised learning algorithm is Convolutional neural network. It consists of number of multiple layers. There are three important layers in CNN they are Convolutional, pooling and fully connected layer. The Convolutional neural network has the advantage of extracting and using features information. For text classification CNN plays a main role in processing structured arrays of data such as images. It vastly used in computer vision and visual application.

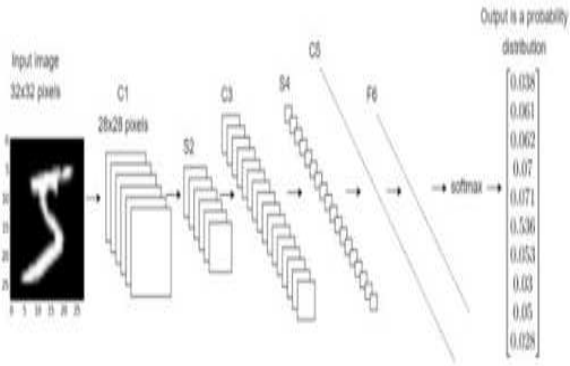


Fig. 3 Block diagram for CNN.

1. Configuration of convolution layer is configured for every set of max pooling layers.
2. To implement lateral inhibition normalization layer is added.
3. The epoch is fixed to 100.
4. The batch size is to 100.

For experimentation purpose sequential model is used to build a CNN model. A learning rate of 0.01 and epsilon is set to le-0.8. loss is as categorical-cross entropy. To find metrics accuracy function is used. These are the python code to configure CNN. The model is differentiated into different layers. Layer 1: A 5× 5 filters is applied to extract sub regions of convolution layer and function tanh is used to activate. Model. add (conv2D (64, (5,5), activation='tanh')).

Layer 2: A pooling layer is configured with size of 5× 5 and 5× 5 strides. Model. Add(max-pooling 2D (poolsize= (5,5), strides (5,5))). In the next step Convolutional layer is configured as said in the above step. A dropout of 0.5 and flatten layer is added for normalization. Model. Add(LRN2D (alpha=0.1, beta=0.75)). Layer 3: A model of number of classes and Softmax function is added to activate. Model. Add (Dense (numofclasses, activation='Softmax')) In the next stage to compile cross entropy categorical is used and optimizer Adam of lr=0.001, beta1=0.9, epsilon of le -08 is set. CNN approach mainly works in two stages namely training and testing, for training the system dataset used is real-time handwritten segmented characters from Kannada words. Data-set contains set of categorical images belong to 100+ classes. Each class consists of 75 handwritten characters. Various pre-processing techniques has been applied that helps to enhance the dataset. The trained images stored in the cnnmodel.h5. Next step is to pass test images pickle and load function is used for test images. A set of test images is passed to compare with trained images, the features of label matches which in evaluates test and train labels. The accuracy is find out by summarize the history at last the loss and epoch is find out and saved in model. Save file.

5 RESULTS AND DISCUSSION

For experimentation purpose, handwritten word recognition dataset is created in such a way that initially questions are formulated from higher school Kannada text book syllabus. Questions is opted from 8th, 9th standard. We have formulated 10 questions that is in the form of printed and for all questions baseline is provided at the end of the question, to write answers above the baseline. Experimentation conducted on different cases i.e. case 1: which contains handwritten answers without ottaksharas and which contains ottaksharas, case 2: Real time data-sets containing noise as these type of dataset is complicated in the case of word recognition. In this way we have created 1000 documents. All the answer script is written by different students from various schools.

CASE 1: Handwritten Answer Scripts containing both question and answer written in a handwritten type.

Answer scripts containing Kannada handwritten answer with and without ottakshara present in a word is collected for experimentation from different schools. The main difficulties of the dataset are that due to non-uniqueness and one or more ottakshara present below the character and drastic variation in handwritten styles.

Here, combination of questions is in the form of printed and baseline is provided to write the answer above the horizontal line and answers is in a handwritten format written by different students, which consists of around 500 scripts of Kannada language. To prove the efficacy of the proposed method, evaluation of answer script with different handwritten answers written by more

Algorithm

Input: Handwritten Kannada Answerscript

Output: Word Recognition

Begin

Step 1: Load the Answerscript.

Step 2: preprocessing is done to remove noise. Dilation and erosion is applied to the image.

Step 3: Histogram is constructed for the binarised image.

Step 4: Horizontal and vertical projection profile is applied to segment.

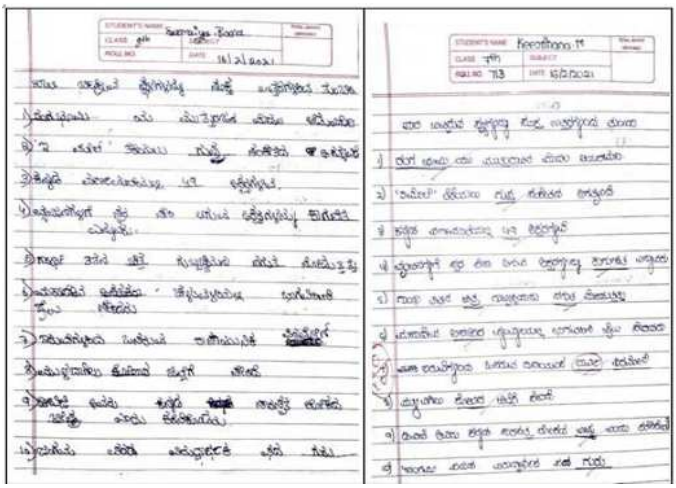
Step 5: Characters is segmented based on x, y, w, h coordinates given using bounding box method.

Step 6: Segmented characters are stored and trained for further recognition stage.

Step 7: Pass CNN to train the characters and test image is passed to compare.

Step 8: Trained image is compared and word is recognized, if the written answer is matched.

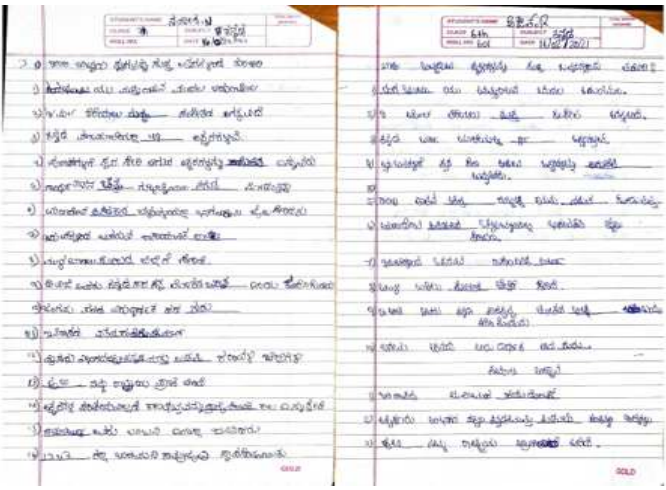
End



(a) (b).

number of students is very big task for teachers to evaluate and to recognize for OCR. Our research work aims at evaluation of one word handwritten answer scripts. Most of the handwritten words are present with ottakshara with one or more touching characters. The extraction of answers from baseline and to segment the character individually is a challenging task. ML and CNN is used for word recognition. The obtained word recognition accuracy is 98.6%.

The bar graph 3 shows that the samples ranging from answer scripts 1to10. The number of answers matched by proposed system. The answers ranging



(c) .

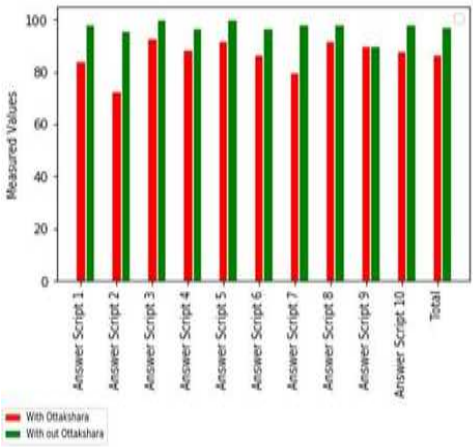


Fig. 4 Graph Representing the Results of Both Question & Answers Written in a Handwritten Type.

from 1,5,7,8 almost coincide with each other which shows high accuracy. For answer script like1,3,4,9 are doesn't match noticeable result. so, there is a difference in number of answer matched.

CASE 2: Match the following with printed questions and handwritten answers.

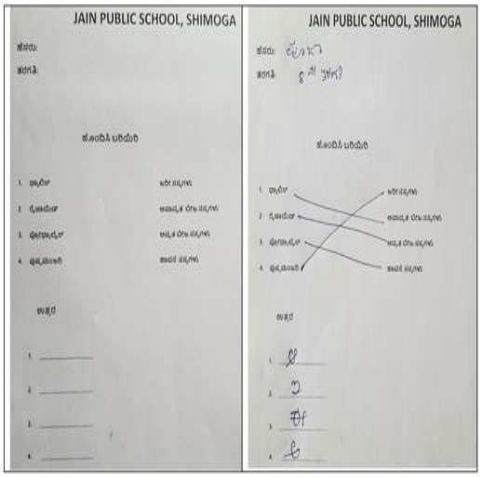


Fig. 5 Sample Answer script of match the following.

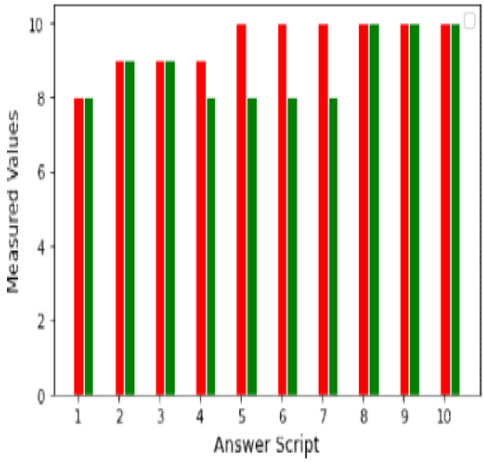
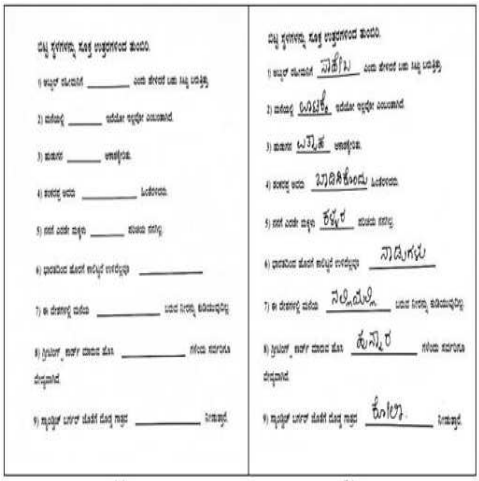


Fig. 6 Representation of Graph depicting the answers for each match the following answers.

The figure 5 shows sample of match the following where around 200 answer scripts is collected from different schools. In the same way as mentioned above for one- word answer. In case of match the following also. In the step1, noise is removed and line detection method is applied to remove baseline and to extract the answer. The extracted answer is stored and trained. In the last step comparison is done with stored answer and recognition is done on the basis of answer matched.

CASE 3: Handwritten Answer Scripts with ottakshara and Without ottakshara.
Answer scripts containing Kannada handwritten answer with and without ottakshara present in a word is collected for experimentation from different schools. The main difficulties of the dataset are that due to non-uniqueness and one or more ottakshara present below the character and drastic variation in handwritten styles.

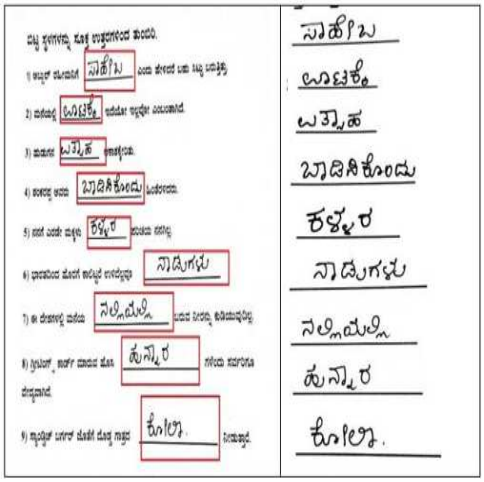


(a) (b).

Here, combination of questions is in the form of printed and baseline is provided to write the answer above the horizontal line and answers is in a handwritten format written by different students, which consists of around 500 scripts of Kannada language. To prove the efficacy of the proposed method, evaluation of answer script with different handwritten answers written by more number of students is very big task for teachers to evaluate and to recognize for OCR. Our research work aims at evaluation of one word handwritten answer scripts. Most of the handwritten words are present with ottakshara with one or more touching characters. The extraction of answers from baseline and to segment the character individually is a challenging task. ML and CNN is used for word recognition. The obtained word recognition accuracy is 98.6%.

The bar graph 7 shows that the samples ranging from answer scripts 1to10. The number of answers matched by proposed system. The answers ranging from 1,5,7,8 almost coincide with each other which shows high accuracy. For answer script like1,3,4,9 are doesn't match noticeable result. so, there is a difference in number of answer matched.

Table 1. Comparison of the Results of proposed method with the existing method.



(c) (d).

(e) (f).

Qualitatively, the results of word Recognition stages are promising. The achieved accuracy represents decent accuracy for a very complex problem, even given for different Kannada handwritten data sets. However, we realize that higher data sizes will make the training phase to be necessary hard to obtain and also increase accuracy significantly, which would undermine the ability to accurately translate handwritten text to a digital form. This led us to begin exploring the word recognition approach.

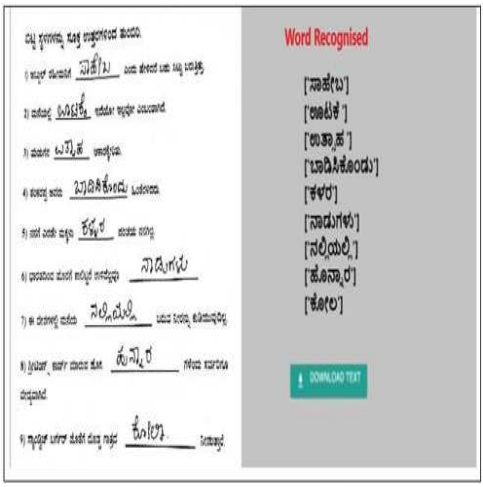


Fig. 7 (g) (h).

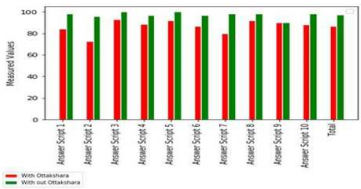


Fig. 8 Graph Representing the Results of Both Question & Answers Written in a Handwritten Type.

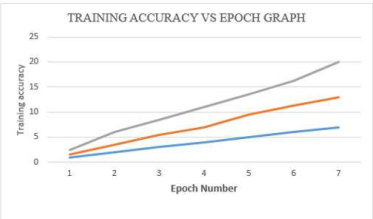
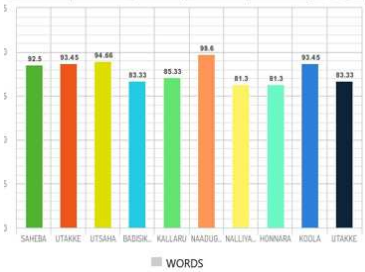


Fig. 9 Graphical representation of training accuracies for each epoch.

Table 2. Recognition Accuracy of Handwritten Kannada Word.

SL.NO	WORDS	NUMBER OF ANSWERS	RECOGNIZED ANSWERS	RECOGNITION RATE	AVERAGE
1	ಸಾಹೇಬ	200	185	92.5%	90.0%
2	ಉತಕೇ	160	150	93.45%	
3	ಉತ್ಸಾಹ	150	142	94.66%	
4	ಬಡಿಸಿಕೊಂಡು	150	125	83.33%	
5	ಕಲ್ಲರು	150	128	85.33%	
6	ನಾಡುಗಳು	150	148	98.6%	
7	ನಲ್ಲಿಯಲ್ಲಿ	150	122	81.3%	
8	ಹೊನ್ನರ	150	122	81.3%	
9	ಕೊಡಲಾ	160	150	93.45%	
10	ಉತಕೇ	150	125	83.33%	



6 CONCLUSION

We have proposed a method to extract answer from scripts and recognition of Kannada answers. The proposed works based on bounding box, straight line detection, feature extraction methods. CNN method is used to train/test recognize handwritten word. Experimentation is carried out on our own dataset containing 1000 one-word answer scripts and proposed approach yields better results for word recognition. The overall accuracy of 96.12% is achieved.

References

[1] C. Naveena, V.N. , Manjunath Aradhya , *Handwritten Character Segmentation for Kannada Scripts* **World Congress on Information and Communication Technologies**, pp. 144-149, 2012.

[2] Mr. Chethan Kumar G.S, Dr. Y.C Kiran, *"Kannada Handwritten Character Segmentation using Curved Line and Recognition using Support Vector Machine"*, **International Journal of Combined Research & Development (IJCRD)**, volume:5, pp. 641-645, 2016.

[3] Mamatha H R, Srikantamurthy K "Morphological Operations and Projection Pro-files based Segmentation of Handwritten Kannada Document" International Journal of Applied Information Systems, Volume 4– No.5, ISSN: 2249-0868, October 2012.

[4] G.Azhaguvarthani, B. Ramani "handwritten text recognition system for English" international journal of pure and applied mathematics, Volume 117 , ISSN: 257-263, No. 16 2017.

[5] RajniKumariSah and K Indira, "Online Kannada Character Recognition using SVM Classifier", in Proceedings of International Conference on Computational Intelligence and Computing Research, 2017.

[6] A. Sushma and Veena.G.S, "Kannada Handwritten Word Conversion to Electronic Textual Format Using HMM Model", in proceedings of International Conference on Computational Systems and Information Systems for Sustainable Solutions 2016.

[7] M.S. Patel, Sanjay Linga Reddy. "An Impact of Grid based Approach in Offline Handwritten Kannada Word Recognition", IEEE-2014.

[8] Aravinda C. V and H.N.Prakash "Template Matching Method for Kannada Hand-written Recognition Based on Correlation Analysis",IEEE-2014.

[9] Saleem Pasha and M.C.Padma, "Recognition of Handwritten Kannada Characters using hybrid Features", in proceedings of IET.

[10] Karthik S, Srikanta Murthy K "Segmentation and Recognition of Handwritten Kannada Text Using Relevance Feedback and Histogram of Oriented Gradients – A Novel Approach" International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, 2016.

[11] A. Alaei, P. Nagabhushan and U. Pal, "A Benchmark Kannada Handwritten Document Dataset and its Segmentation", International Conference on Document Analysis and Recognition, pp. 141-145, 2011.

[12] Tulika Sureka , Swetha K S N , Isha Arora and Dr. Mamatha H R "Word Recognition Techniques For Kannada Handwritten Documents" IEEE,10th ICCCNT 2019 July 6 -8, 2019 - IIT, Kanpur, Kanpur, India.

[13] https://en.wikipedia.org/wiki/Kannada.

[14] Shakunthala B S Naveen N "Unconstrained Handwritten Kannada Documents leading to Line and Word segmentation" International Journal of Engineering Re-search Technology, 2018.

[15] Subhrajyoti Sen, Shreya V Prabhu, Steve Jerold, Pradeep J S, Dr. Savita Choudhary" Comparative Study and Implementation of Supervised and Unsupervised Models for Recognizing Handwritten Kannada Characters", 3rd IEEE Conference on Recent Trends in Electronics, Information International Communication Technolo-gy (RTEICT-2018), pp.774-778,2018.

[16] Chollet, Francois and others, "Keras", https://keras.io, 2015.

[17] Leslie N. Smith, "Cyclical Learning Rates for Training Neural Networks", WACV 2017.

[18] Andreas C. Muller, Sarah Guido, "Introduction to Machine Learning with Python - A guide for Data Scientists", Sebastopol, CA, O'ReillyMedia, Inc. 2016.

[19] Ian Goodfellow, Yoshua Bengio, Aaron Courville, "Deep Learning", MIT Press, 3 January 2017.

[20] Denise Klep, Gerven Sanne Schoen makers, Marcel Van, "Data augmentation of a handwritten character dataset for a Convolutional Neural Network and integration into a Bayesian Linear Framework", June 3, 2016.

[21] Ms. Paden Rinchen, "Comparative Study of Techniques used for Automatic Evaluation of Free Text Answer", International Journal on Recent and Innovation Trends in Computing and Communication. Pp no 3839-3842, 2014.

[22] Divya Patel1, Shaikhji Zaid2, "Efficient System for Evaluation of Omr Sheet a Survey", international journal of advanced research in engineering, science and management, Pp no01- 08, 2017.

[23] K. Ntirogiannis, B. Gatos and I. Pratikakis, "An Objective Evaluation Methodology for Handwritten Image Document Binarization Techniques", Pp.217-224, 2008.

[24] NilimaSandipGite, “Implementation of Descriptive Examination and Assessment System”, international journal of advance research in science and engineering, volume no.07, Pp.252-257, 2018.

[25] VaishaliMalpe, Shalini Bhatia, “Evaluation of Students’ Answer Scripts Using Soft Computing Techniques”, International Journal of Modern Engineering Research (IJMER), Vol.2, Issue.3, pp-1280-1289, 2012.

[26] TilaniGunawardena, NisharaPathirana, MedhaviLokuhetti, Roshan Ragel, and SampathDeegalla, “Performance Evaluation Techniques for an Automatic Question Answering System”, International Journal of Machine Learning and Computing, Vol. 5, No. 4, Pp.294-300,2015.

[27] Andrea Andrenucci, EriksSneiders, “Automated Question Answering: Review of the Main Approaches”, Proceedings of the Third International Conference on In- formation Technology and Applications.Pp.514-519, 2005.

[28] TilaniGunawardena, MedhaviLokuhetti, NisharaPathirana, Roshan Ragel and SampathDeegalla “An Automatic Answering System with Template Matching for Natural Language Questions”, IEEE, Pp.353-358, 2010.

[29] S. Prayag Singh, SaurabhSheorain, ShivamTomar, Shubham Sharma, N.K.Bansode , “Descriptive Answer Evaluation”, International Research Journal of Engineering and technology, Volume: 05, Pp.2709-2712, 2018.